SCHOLARLY PUBLICATIONS

A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT

@DTU

(102nd Edition)

JUNE 2021

BY: CENTRAL LIBRARY

DELHI TECHNOLOGICAL UNIVERSITY
(FORMERLY *DELHI COLLEGE OF ENGINEERING)*
GOVT. OF N.C.T. OF DELHI
SHAHBAD DAULATPUR, MAIN BAWANA ROAD
DELHI 110042

PREFACE

This is the **Hundred Second** Issue of Current Awareness Bulletin started by Delhi Technological University, Central Library. The aim of the bulletin is to compile, preserve and disseminate information published by the faculty, students and alumni for mutual benefits. The bulletin also aims to propagate the intellectual contribution of Delhi Technological University (DTU) as a whole to the academia.

The bulletin contains information resources available in the internet in the form of articles, reports, presentations published in international journals, websites, etc. by the faculty and students of DTU. The publications of faculty and student which are not covered in this bulletin may be because of the reason that the full text either was not accessible or could not be searched by the search engine used by the library for this purpose.

The learned faculty and students are requested to provide their uncovered publications to the library either through email or in CD, etc. to make the bulletin more comprehensive.

This issue contains the information published during **June, 2021**. The arrangement of the contents is alphabetical. The full text of the article which is either subscribed by the university or available in the web is provided in this bulletin.

Central Library

CONTENTS

1. 3D skeletal gesture recognition through discriminative coding, ***6.Barkha singh*** and ***3.Prof.J.Panda***, ECE, DTU

2. A 32 nm single ended single port 7T SRAM for low power utilization, ***6.Bhawna Rawat*** and ***3.Poornima Mittal***, Electronics, DTU

3. A first principles based study of the effect of uniform and tetragonal strains on half-metallicity in FeCrAs Heusler alloy, ***8.Aniket Singh***, ***8.Mansher Singh*** and ***3.Mukhtiyar Singh***, Physics, DTU

4. A Literature Review on Brain Tumor Detection and Segmentation, ***8.Aditya Miglani***, ***8.Hrithik Madan***, ***8.Saurabh Kumar*** and ***3.Sanjay Kumar***, CSE, DTU

5. A Novel 8.5 GHz Parallel Coupled, Half Wavelength Bandpass Filter with Dumbbell-Shaped Defected Ground Structure, ***8.Aditya Agrawal***, ***8.Adya Agarwal***, ***8.Ojas Aryam*** and ***3.Asok De***, Electronics, DTU

6. A novel quantum inspired hybrid metaheuristic for dispatch of power system including solar photovoltaic generation, Kapil Deo Bodha, ***3.Vinod Kumar Yadav*** and Vivekananda Mukherjeea, Electrical, DTU

7. A simple and novel SAW magnetic sensor with PVA boundmagnetostrictive nanopowder film, Fahima, ***7.1.Mainuddin***, Pooja Rajput, Jitender Kumar and A.T. Nimal, Electronics, DTU

8. A Study of thermoelectric properties of Hf-doped RhTiP Half-Heusler alloy, ***6.Shivam Tyagi***, ***7.Shivani Arora*** and ***3.Mukhtiyar Singh***, Physics, DTU

9. A Survey on Artificial Neural Network based Stock Price Prediction Using Various Methods, ***3.Shailender Kumar***, ***8.Rounak Gupta***, ***8.Puneet Kumar*** and ***8.Naman Aggarwal***, CSE, DTU

10. Adaptive control using stochastic approach for unknown but bounded disturbances and its application in balancing control, ***6.Rupam Singh*** and ***3.Bharat Bhushan***, Electrical, DTU

11. Adaptive Histogram Equalization based Approach for SAR Image Enhancement: A Comparative analysis, ***8.Vinay Dubey*** and ***3.Rahul Katarya***, CSE, DTU

12. ADAS Authentic Data Allowed Security In Internet Of Things, ***6.Raveena Yadav*** and ***3.Vinod Kumar***, CSE, DTU

13. AI based Impact of COVID 19 on food industry and technological approach to mitigate, ***3.Ram Murti Rawat***, ***8.Ankit Jose Toppo***, ***8.Ashish Rana*** and ***8.Akshit Beck***, CSE, DTU

14. An FEA based study of thermal behaviour of ultrasonically welded phosphor bronze sheets, Bharat Sanga, ***3.Reeta Wattal*** and ***3.D.S. Nagesh***, Mechanical, DTU

15. An Introduction to Signed Petri Net, ***3.Sangita Kansal*** and ***6.Payal Dabas***, Mathematics, DTU

16. Analytical modeling of dual-metal gate stack engineered junctionless accumulation-mode cylindrical surrounding gate (DMGSE-JAM-CSG) MOSFET, ***8.1.Sumedha Gupta***, ***3.Neeta Pandey*** and R. S. Gupta, Electronics, DTU

17. Assessment of phytochemical and genetic diversity analysis of Plumbago zeylanica L. accessions, ***6.Arpita Roy***, Neelam Sharma and ***3.Navneeta Bharadvaja***, Biotechnology, DTU

18. BLOCKCHAIN TECHNOLOGY IN PREVENTING THE THEFT OF HUMAN ORGAN AND RARE BLOOD GROUP, ***7.1.Yogesh Sharma*** and B. Balamurugan, Computer, DTU

19. Capacitive Analysis of a Novel 3 nm Truncated Fin FinFET, ***6.Mridul Prakash Kashyap*** and ***3.Rishu Chaujar***, Electronics, DTU

20. CFOA Based Negative Floating Fractional Order Capacitance Multiplier, ***8.Rahul Gupta***, ***8.Charvi Agrawal***, ***8.Ankit Gupta*** and ***3.Rajeshwari Pandey***, Electronics, DTU

21. Characterization of on Newly Developed Hybrid Reinforced Composite of Polypropylene with Agave –Amaricana, Nano-talc and Starch, G S Bajwa, ***3.Ranganath M Singari*** and ***3.R S Mishra***, Mechanical, DTU

22. Classification of ancient inscription images on the basis of material of the inscriptions, ***3.N. Jayanthi***, ***8.Tarush Sharma***, ***8.Vinay Sharma***, ***8.Shubham Tyagi*** and ***3.S. Indu***, ECE, DTU

23. Classification of Human Activities using data captured through a smartphone using deep learning techniques, ***7.Lokesh Dhammi*** and ***3.Piyush Tewari***, Electronics, DTU

24. CNN-based Person Recognition System for Masked Faces in a post-pandemic world, ***8.Parangat Mittal***, ***8.Kartikeya Pandey***, ***8.1.Peeyush Tawani*** and ***3.Prof. Rajesh Rohilla***, Electroncis, DTU

25. Comparative Analysis of the Design Techniques for Low Leakage SRAMs at 32nm, ***7.Monica Gupta***, Kirti Gupta and ***3.Neeta Pandey***, Electronics, DTU

26. Comparative Study of Feature Reduction Techniques in Software Change rediction, ***3.Ruchika Malhotra***, ***8.Ritivik Kapoor***, ***8.Deepti Aggarwal*** and ***8.Priya Garg***, Sofware Engineering, DTU

27. Consensus Sequence Characterization Of Rnp 1 And Rnp 2 Along With Base Stacking Interactions With Ribonucleuotide, ***7.Isha Sehrawat*** & ***3.Yasha Hasija***, Biotechnology, DTU

28. COVID 19, Pneumonia and Other Disease Classification using Chest X-Ray images, ***8.Abhishek Panwar***, ***8.Akshay Dagar***, ***8.Vishal Pal*** and ***3.Dr. Vinod Kumar***, CSE, DTU

29. COVID-19 Severity Assessment from Chest X-rays using Attention-based Weakly-Supervised Learning, ***8.Kunal Chaturvedi***, ***8.Tanmay Kansal***, ***8.Saksham Gupta***, ***3.Dinesh Kumar Vishwakarma*** and ***3.Naokant Deo***, Mathematics and IT, DTU

30. Design and Study of Dadda Multiplier by using 4:2 Compressors and Parallel Prefix Adders for VLSI Circuit Designs, ***8.Ayush Kumar***, ***8.Ankit Kumar*** and ***3.Dr. Deva Nand***, ECE, DTU

31. Design Of High Bandwidth Circularly Polarised Antipodal Vivaldi Array for 5G Applications, ***8.Anirudh Nakra***, ***8.Abhijeet Vats*** and ***3.Asok De***, Electronics, DTU

32. Detection of Pulsars by Classical Machine Learning Algorithms, ***8.Deepanshu Beniwal***, ***8.Abhishek Roy***, ***8.Himanshu Yadav*** and ***3.Anamika Chauhan***, IT, DTU

33. Detection Of Skin Cancer Using Neural Architecture Search with Model Quantization, ***8.Ritabrata Maiti***, ***8.Priyanka Agarwal***, ***8.Ragyayee Ravinder Kumar*** and ***3.Aruna Bhat***, CSE, DTU

34. Determination of the Probabilities of Landslide Events in Kalimpong, West Bengal, ***8.Parvesh Lamba***, ***8.Pooja***, ***8.Amit Dabas***, ***8.Hitesh Khatri*** and ***3.Prof. S. Anbu Kumar***, Civil, DTU

35. Dielectric Engineered Schottky Barrier MOSFET for Biosensor Applications: Proposal and Investigation, ***8.1.Rahul Singh***, ***8.Shweta Kaim***, ***6.Rani Medha Shree***, ***8.Anil Kumar*** and ***3.Sumit Kale***, Electronics, DTU

36. Embodying Current division transconductance reduction in M-CDTA and its application as low frequency filter, ***8.Mohit Bansal***, ***8.Pranav Semwal***, ***8.Navneet Sagar*** and ***3.Neeta Pandey***, Electroncis, DTU

37. Evaluation of Different Phytochemicals Against BRCA2 Receptor, Karanpreet Singh Bhatia, ***8.Saksham Garg***, ***8.Ashutosh Anand*** and Arpita Roy, DTU

38. Fabrication and multi-objective optimization of friction stir processed aluminium based surface composites using Taguchi approach, ***3.Ravi Butola***, ***3.Prakash Chandra***, ***8.Kartikeya Bector*** and ***3.RanganathMSingari***, Mechanical, and Biotechnology, DTU

39. Fake Currency Detection with Machine Learning Algorithm and Image Processing, ***8.Aman Bhatia***, ***8.Vansh Kedia***, ***8.Anshul Shroff***, ***8.Mayand Kumar***, ***8.Bickey Kumar Shah*** and Aryan, CSE, DTU

40. Feature Extraction Techniques in Facial Expression Recognition, ***7.Nitendra Mishra*** and ***3.Aruna Bhatt***, CSE, DTU

41. FPGA Implementation of a pipelined and pseudo-randomized TDES algorithm, ***8.Arnab Roy***, ***8.Anubhav Gadhok*** and ***3.Deva Nand***, Electronics, DTU

42. From Classical to Quantum: A Review of Recent Progress in Reinforcement Learning, ***8.Kunal Kashyap***, ***8.Lalit***, ***8.Daksh Shah*** and ***3.Lokesh Gautam***, Electronics, DTU

43. Gain Enhancement Using Modified Circular Loop FSS Loaded with Slot Antenna for Sub-6GHz 5G Application, Anubhav Kumar, ***3.Asok De*** and Rakesh K. Jain, Electronics, DTU

44. Generating Caption for Image using Beam Search and Analyzation with Unsupervised Image Captioning Algorithm, ***3.Prashant Giridhar Shambharkar***, ***8.Priyanka Kumari***, ***8.Pratik Yadav*** and ***8.Rajat Kumar***, CSE, DTU

45. Genre Classification using Character Networks, ***3.Rahul, 8.Ayush***, ***8.Divya Agarwal*** and ***8.Devika Vijay***, CSE, DTU

46. Green Synthesis of Nanoparticles and its Application in Sewage Treatment, Mikhlesh Kumari, Sunita Punia, ***6.Rajesh Kumar Lohchab*** and ***3.Anil Kumar Haritash***, Environmental, DTU

47. Heat and water flux modeling in an earth dam, ***6.Sushant Kumar***, ***6.A. K. Sahu*** and ***3.Munendra Kumar***, Civil, DTU

48. High Performance Multipliers Using Hybrid Adder, ***8.Sumit Raj***, ***8.Utkarsh Chaurasia***, ***8.Aayush Bahukhandi*** and ***3.Poornima Mittal***, Electronics, DTU

49. How Low is Too Low? A Computational Perspective on Extremely Low-Resource Languages, ***8.Rachit Bansal***, ***8.1.Himanshu Choudhary***, ***8.1.Ravneet Punia***, Niko Schenk, Jacob L Dahl and E´milie Page´-Perron, CSE, DTU

50. Hydrogen Production from Biogas: Methods and Economic Analysis, ***6.Ravindra Kumar***, ***6.Anil Kumar*** and ***3.Amit Pal***, Mechanical, DTU

51. IAF Transition towards a Network-Centric- Knowledge Based Force, Wing Commander, Mukesh Kumar Singh and ***3.Dr Vikas Gupta***, DSM, DTU

52. Identifying influential nodes in weighted complex networks using an improvedWVoteRank approach, ***3.Sanjay Kumar*** and Ankit Panda, CSE, DTU

53. Image Detection in Noisy Images, ***8.Kushagra Yadav***, ***8.Dakshit Mohan*** and ***3.Anil Singh Parihar***, CSE, DTU

54. Impact of interface trap charges on Junctionless double and triple metal gate High-k Gate All Around Nanowire FET based Alzheimer Biosensor, ***3.Rishu Chaujar*** and ***6.Mekonnen Getnet Yirak***, Electronics, DTU

55. Implementation of a Shadow Filter using CFCC, ***8.Prakhar Prakash***, ***8.Yashodhan Srivastava***, ***8.Shubham Kumar*** and ***3.Alok Kumar Singh***, Electronics, DTU

56. Importance of Cytochrome P450 gene family from metabolite biosynthesis to stress tolerance: A review, ***6.N Laffaru Singpho*** and ***3.J G Sharma***, Biotechnology, DTU

57. Insights and reflections using data analytics approach of COVID- 19 situation among the Indian states and ethical considerations, Zairu Nisha, Sachin Kumar, Saibal K. Pal, ***3.Jagvinder Singh*** and Ompal Singh, USME, DTU

58. Intra-Block Correlation Based Reversible Data Hiding in Encrypted Images Using Parametric Binary Tree Labeling, Arun Kumar Rai, Neeraj Kumar, ***3.Rajeev Kumar***, Hari Om, Satish Chand and Ki-Hyun Jung, CSE, DTU

59. Isolation, Synthesis, and Medicinal applications of Heparin, ***7.Saikrushna Jena*** and ***3.Ram Singh***, Chemistry, DTU

60. Life dissatisfaction among students: exploring the role of intrapersonal conflict, insufficient efforts and academic stress, Anju, Amandeep, B.K. Punia, Vandana Punia and ***3.Naval Garg***, DTU

61. Lipase catalyzed esterification of Docosahexaenoic acid (DHA) with immobilized Pseudomonas cepacia and Thermomyces lanuginosus, Aditi Sharma, ***3.Kriti Bhandari***, Anjali Jain, Satyendra P. Chaurasia, and Ajay K. Dalai, Biotechnology, DTU

62. Local Moment driven PVO based Reversible Data Hiding, Neeraj Kumar, ***3.Rajeev Kumar*** and Roberto Caldelli, CSE, DTU

63. Lower Fin Modulation Analysis for a Novel 5nm Top Bottom Gated Junctionless FinFET for improved performance, ***8.Lakshya Gangwani***, ***8.Ruchika Chakravarti*** and ***3.Rishu Chaujar***, Physics, DTU

64. LRD-Net: A Lightweight Deep Network for Low-light Image Enhancement, ***8.Shivam Singhal***, ***8.Srishti Nanduri***, ***8.Yash Raghav*** and ***3.Anil Singh Parihar***, CSE, DTU

65. Machine Learning and IoT based solutions for detection of arrhythmia using ECG signals, ***3.Vinod Kumar***, ***8.Awishkar Ghimire*** and ***8.Hitesh Kumar Hoon***, CSE, DTU

66. MDER:modified degree with exclusion ratio algorithm for influence maximisation in social networks, ***3.Sanjay Kumar***, ***8.Dipti Lohia***, ***8.Darsh Pratap***, ***8.Ashutosh Krishna*** and B. S. Panda, CSE, DTU

67. Medical Image Fusion using Lifting based Transform Techniques, ***8.Anisha Gupta***, ***8.Priyesh Chaurasia*** and ***3.Sudipta Majumdar***, Electronics, DTU

68. Microbial Amelioration of Salinity Stress in HD 2967 Wheat Cultivar by Up-Regulating Antioxidant Defense, Madhulika Singha and ***6.Neha Tiwari***, Biotechnology, DTU

69. Movie genres and beyond: An analytical survey of classification techniques, ***8.Avleen Singh Mehal***, ***8.Kalpana Meena***, ***8.Rijul Bir Singh*** and ***3.Prashant Giridhar Shambharkar***, CSE and Software, DTU

70. Multibeam antenna using cavity-based waveguide structure with extended hemispherical lens at millimetre wave, Shakti Singh Chauhan, ***7.1.Harikesh Dalal***, Mahesh P. Abegaonkar and Ananjan Basu, ECE, DTU

71. Multiplex Regulation System With Personalised Recommendation Using ML, ***8.Anshul Shroff***, ***8.Bickey Kumar Shah***, ***8.Aayush Jha***, ***Amar Kumar 8.Jaiswal***, ***Pooja Sapra*** and ***3.Manoj Kumar***, CSE and Civil, DTU

72. Natural Microalgae Biofilm As A Tool To Clean-Up Mining Wastewater, ***3.S. Anbu Kumar***, ***8.Shivam Gupta***, ***8.Shubham Kumar*** and ***8.Pawan Kumar Meena***, Civil, DTU

73. Object Tracking in a Zone using DeepSORT, YOLOv4 and TensorFlow, ***3.Dr. Rajni Jindal***, ***8.Aditya Panwar***, ***8.Nishant Sharma*** and ***8.Aman Rai***, CSE, DTU

74. ON CERTAIN GENERALIZATIONS OF S————————————( ): II, ***6.KAMALJEET GANGANIA*** & ***3.S. SIVAPRASAD KUMAR***, Mathematics, DTU

75. Penta -hexa coordination behaviour of ABA-P(V) corrole, ***6.Ritika Kubba***, Omprakash Yadav, Pinky Yadav, Natalia Fridman and ***3.Anil Kumar***, Chemistry, DTU

76. Performance Analysis of PV Arrays under Different Shading Conditions, ***8.Abhishek Yadav***, ***8.Aaradhya Mittal***, ***8.Nadeem Hussain Wazir***, ***3.Vinod Kumar Yadav*** and Santosh Ghosh, Electrical, DTU

77. Plasmonic behaviour of a 1-D plasmonic nanostructure on a spacer-separated metallic film, ***7.Sreelekshmi R. Nair*** and ***3.Yashna Sharma***, Electronics, DTU

78. QnA System on Educational Textbooks: Digital Library Doubt Support System, ***3.Rahul Gupta***, ***8.Gaurav Dabas***, ***8.Himanshu Yadav*** and ***8.Nijdat Hasnain***, CSE, DTU

79. Relevance of phosphate solubilizing microbes in sustainable crop production: a review, F. Fatima, M. M. Ahmad, ***3.S. R. Verma*** & N. Pathak, Biotechnology, DTU

80. Revisiting the Factors Influencing Teaching Choice Framework: Exploring What Fits with Virtual Teaching, Sudhir Rana, Amit Kumar Singh, ***6.Shubham Singhania***, Shubhangi Verma and Moon Moon Haque, USME, DTU

81. SAGA: A Hybrid Technique to handle Imbalance Data in Software Defect Prediction, ***3.Ruchika Malhotra***, ***8.Ritvik Kapoor***, ***8.Paridhi Saxena*** and ***8.Parth Sharma***, CSE, DTU

82. Scenario of Biomedical Waste Management during COVID-19 Pandemic in Delhi, India, ***8.Divyanshu Aneja***, ***8.Anshul Rana***, ***8.Anjali Kumari*** and ***3.Anunay A. Gour***, Environment, DTU

83. Sequential Recommendation with Temporal Context via Convolutional Sequence Embedding, ***3.Dr. Aruna Bhat***, ***8.Rishabh Chandra*** and ***8.Kanika***, CSE, DTU

84. SIGN LANGUAGE ALPHABET RECOGNITION USING CONVOLUTION NEURAL NETWORK, ***8.Mayand Kumar***, ***8.Piyush Gupta***, ***8.Rahul Kumar Jha***, ***8.Aman Bhatia***, ***8.Khushi Jha*** and ***8.Bickey Kumar Shah***, Software and CSE, DTU

85. Significant improvements in dielectric, impedance, multiferroic and magnetoelectric properties of (1 2 x)Co0.5Ni0.5Fe2O42xBaTiO3 bulk composites (x = 0, 0.10 and 0.20), ***6.S. Shankar***, O. P. Thakur and ***3.M. Jayasimhadri***, Physics, DTU

86. Solar Energy Prediction using Decision Tree Regressor, ***3.Rahul***, ***8.Aakash Gupta***, ***8.Ankur Bansal*** and ***8.Kshitij Roy***, CSE, DTU

87. Strong enhancement in structural, dielectric, impedance and magnetoelectric properties of NdMnO3 - BaTiO3 multiferroic composites, ***6.S. Shankar***, O.P. Thakur and ***3.M. Jayasimhadri***, Physics, DTU

88. Structural and photoluminescence properties of Er3+ doped SrBi2Nb2O9 ceramics, ***7.Ritushree Shaily*** and ***3.Renuka Bokolia***, Physics, DTU

89. Structural properties of Strontium Bismuth Niobate (SrBi2Nb2O9) ferroelectric ceramics, ***7.Pooja Mojumdar***, ***7.Ritushree Shaily*** and ***3.Renuka Bokolia***, Physics, DTU

90. Study of residual stresses in multi-pass friction stir processed surface composites, ***8.Kartikeya Bector***, ***8.Mrinal Singh***, ***8.Divya Pandey***, ***3.Ravi Butola*** and ***3.Ranganath M. Singari***, Mechanical, DTU

91. Study of robust facial recognition under occlusion using different techniques, ***7.Deepanshu Jayant*** and ***3.Dr. Aruna Bhat***, CSE, DTU

92. Super-Capacitor based Electric Vehicle Electric Vehicle Charging, ***8.Hemant Sharma***, ***6.Imroz Khan***, ***8.Anjali Kumari*** and ***3.Prof. Prem Prakash***, Electrical, DTU

93. TCAD Analysis and Simulation of Double Metal Negative Capacitance FET (DM NCFET), ***6.Yash Pathak***, ***3.Bansi Dhar Malhotra*** and ***3.Rishu Chaujar***, Physics and Biotechnology, DTU

1. *Vice Chancellor*
2. *Pro Vice Chancellor*
3. *Faculty*
4. *Teaching-cum-Research Fellow*
5. *Asst. Librarian*
6. *Research Scholar*
7. *PG Scholar*
8. *Undergraduate Student*

1.1. *Ex Vice chancellor*
2.1. *Ex Pro Vice Chancellor*
3.1. *Ex Faculty*
4.1. *Alumni*
5.1 *Others*
6.1. *Ex Research Scholar*
7.1. *Ex PG Scholar*
8.1. *Ex Undergraduate Student*

Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021)
IEEE Xplore Part Number: CFP21K74-ART; ISBN: 978-0-7381-1327-2

3D skeletal gesture recognition through discriminative coding

Human activity recognition using time warping invariant Riemannian trajectories

Barkha singh
ECE Department
Delhi Technological University
Delhi, India
Singh.barkha718@gmail.com

Prof.J.Panda
ECE Department
Delhi Technological University
Delhi, India
jpanada@dce.ac.in

2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) | 978-1-6654-1272-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICCS51141.2021.9432304

***Abstract*—To recognize the human activity is considered very essential in human-to-human interaction and interpersonal relations due to its nature of providing information regarding the identity of a people, their individuality, and psychological state. The extraction of this information is very challenging. The extraction of this information is very challenging. The major subject of study of the scientific areas of computer vision and ML is the potential of human for identifying the activities of another person. A sequence of human body movements in which different body parts are engaged in a concurrent manner is known as action. According to computer vision perception, any kind of observation is matched with earlier defined patterns and then it is labeled while recognizing the action. In this research work, the 3 D Skelton-based technique is proposed for human activity reorganization. The proposed technique will improve performance for the human activity detection in terms of accuracy, precision and recall also the proposed technique will compared with the other techniques which authenticates reliability of the model.**

Keywords—***Human activity recognition (HAR) Dynamic time warping (DTW)Convolutional neural network (CNN) Recurrent neural network (RNN) Hidden Markov model (HMM) Dense Trajectories (DT) Motion energy image (MEI) Motion history image (MHI) spatial temporal interest point (STIP) probabilistic regression Neural Network (PRNN) Autoencoder PRNN (APRNN) support vector machine (SVM) Histogram of Oriented Gradients(HOG) Decision through Holistic Intelligence (DHI)***

I. INTRODUCTION

Every action, even a minor action of a human is performed with some purpose. The major subject of study of the scientific areas and ML is the potential of humans for identifying the activities of another person [1]. The classification of human actions is done in four levels on the basis of complexity. Various components of these kinds of systems are focused on extracting features, learning and classifying the action, and also on recognizing and segmenting the action. There are three stages included in a simple procedure in which the human body is detected, tracked and the tracking results are utilized to execute recognition. For instance, the arms and hands of two individuals are detected and tracked to recognize the activities of shaking hands so that a spatial-temporal description of their movement is generated. Comparison of this description is done with the existing patterns in the training data so that the action type is recognized. Stages in Human Action Recognition: In general, a human action recognition framework consists of three main stages.

I**. Segmentation**: Initially, the estimation of background is taken place to be the preceding frame. Afterward, the elimination of this background is done from the input frame. The foreground mask is achieved employing a threshold to the absolute difference [2]. This approach is helpful or not has determined in terms of speed, frame rate, and global threshold.

ii. **Feature selection and extraction**: This process is aimed at achieving the attributes of activity and analyzing them later on. An imperative design decision is to build the feature vector selecting the features in the subsequent module for classifying the features of the whole solution.

iii. **Feature classification**: Several ML algorithms are be implemented for classifying the features that are extracted in the preceding stage when the feature extraction is completed and those features are selected [3]. The overall sketch of these stages is represented with a block diagram expressed in Fig.1

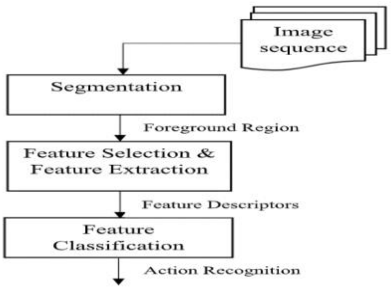


Figure 1: A General Human Action Recognition Framework

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:45:36 UTC from IEEE Xplore. Restrictions apply.

Literature Review

Xinyi Liang, et.al [8] discussed the novel Skeleton-based action identification in which joint trajectory images and vision features were utilized for modeling human action.

Qinkun Xiao, et.al [9] designed a DNN framework for recognizing human action.

Thien Huynh-The, et.al [10] recommended a new hierarchical deep feature fusion model to recognize the human action based on a three-dimensional skeleton.

B Jagadeesh, et.al [11] Initially, did the extraction of100 frames was from every video sequence and the optical flow evaluation among the frames.

Amir Nadeem, et.al [12] focused on incorporating the linear discriminant analysis with ANN in order to recognize and detect the exact human action.

Yun Han, et.al [13] suggested a GSA framework in order to model the weights for skeletal joints and the notion of the ALC model for gathering the frames which played a significant role to make the decision.

Sharmin Majumder, et.al [14] stated that the implementation of HAR in a number of applications.

Chen Chen, et.al [15] intended the HAR system. The overall classification rate obtained from the devised system was evaluated above 97.

Hang Yan, et.al [16] introduced an effective approach to identify the continuous human rehabilitation action in real-time.

Naresh Kumar, et.al [17] emphasized HAR based on skeleton and presented a motion trajectory computation method for which Fourier temporal attributes taken from the interpolation of skeleton joints of the human body were utilized.

Suraj Prakash Sahoo, et.al [18] designed a DBiLSTM for sequential learning while establishing the temporal association among the action frames.

Carlos Roig, et.al [19] suggested a new multi-modal technique in which task-specific attributes were incorporated in order to identify human action.

Xueping Liu, et.al [20] focused on exploiting the Kinect extraction and human activity recognition innovation.

Meng-Lin Chiang, et.al established a vision-based HAR system for supporting human interaction for companion robots [21].

Tin ZarWint Cho, et.al [22] intended a system that emphasized improving HAR with the implementation of skeletal features taken from Kinect sensor so that the discriminative features were achieved

Peng Wang, et.al [23] recommended a HAR algorithm planed on the basis of CNN whose execution was done for identifying the human semaphore motion.

Sui Longfei, et.al [24]suggested a technique for recognizing the human body action with three-dimensional data of human skeleton acquired deploying Kinect depth camera.

Sheng Li, et.al introduced a compression algorithm named MJS[25]. The compression was carried out on the basis of the amount of movement of different joints in this algorithm

Chengwu Liang, et.al [26] intended a segmental framework so that the relation of sub-actions was associated with heterogeneous information fusion and CPPCR for multi-modal HAR [26].

Following are the various conclude points after review: -

1. The techniques which are proposed previously extract a smaller number of features for human activity detection which leads to low accuracy of the system.
2. In human activity detection, classification techniques are used in the system. The complexity of the systems is maximized due to the classification methods.

II. PROPOSE METHOD

HAR (Human Action Recognition) is generally categorized into two categories of action classification and action detection. In action classification, a segmented video comprising just a single action is analyzed. This action needs to be classified into a clear action category Action detection, on the other hand, detects the beginning and end times of all actions in video footage, traces their position in space, and detects the category of action.

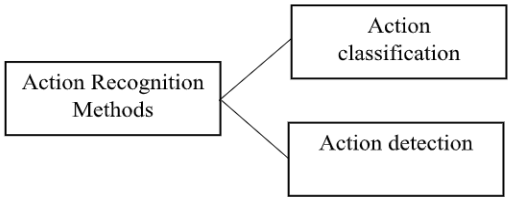


Figure 2: Classification Schema for Human Action Recognition Techniques

A. Action classification:

In the classification stage of HAR, the activities signified by appropriate feature sets extracted from images or videos are

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:45:36 UTC from IEEE Xplore. Restrictions apply.

classified. In this phase, the activity label is provided as final result by the classification algorithm. In fact, almost all classification techniques can be categorized into 3 types of models. All these methods have been described below:

1. **Template-Based Approaches**: This approach focuses on representing general appearance features of a particular action with the help of several depictions. These general appearance features, like two-dimensional or three-dimensional static images/volumes or a series of view models, are called templates. Almost all template-based techniques after extracting two-dimensional or three-dimensional static templates make a comparison of the retrieved images/volumes of test videos and the stored templates on the basis of similarity [4]. Dynamic Time Warping (DTW) is a competent method to perform classification on the basis of a series of keyframes.

One of popular template-based approach is:

Template Matching: These method employs 3-D spatiotemporal template to recognize the actions. These techniques aim at constructing a realistic action template and carry out matching process efficiently according to this template. It is the former method to HAR. A well-known template matching method uses MEI and MHI to represent action. MEI and MHI represent the occurrence of motion and the regency of motion respectively. These are generated for all templates of an action. Actually, the template images, which are created, considered as weighted projection of the space-time shape

2.STIP-based methods: In these methods, the main area of movement variation is extracted from a video for action representation. Like the local features for object recognition in images, it is essential for these methods to make a decision about the appropriate method of key region detection, feature vector to define the key region, and the classification algorithm to use. In STIP, the "interesting points" usually correspond to position changing most vividly in the spatiotemporal dimension. Some well-known STIP techniques are three-dimensional-Harris ST feature points and enhanced methods.

3. **Dynamic Time Warping**: DTW is a type of dynamic programming algorithm that match two sequences with variance. This approach was mainly designed to recognize the speech, indicating the words as template sequence, and give matching scores for fresh words. This method can be applied to recognize the activities of humans as human actions may be observed as a series of keyframes. The conversion of recognition issue has been carried out for the template matching task. The gestures are represented by a set having learned view models. After that, the gesture template is matched amid image frames and view models using the DTW algorithm [5]. Though a few numbers of training samples are required in this algorithm, the computation becomes quite difficult to handle the maximization of action types as these invariances can be stored with the help of extensive templates.

Generative Models: There are various kinds of generative models that are defined as:

1. **Hidden Markov Model Approach**: It is a critical evaluation issue that is one among the three problems of HMM and the forward algorithm is implemented to tackle this issue. This model is suggested for dealing with issue of recognizing the speech [8]. First of all, this model was deployed in the identification of actions. Features, which reveal the pixels in every divided mesh, can be acquired as observations for every frame. Afterward, the observation feature vector sequences are utilized for every action in order to train the Hidden Markov Model in which the preliminary probability of concealed positions is contained. The depiction that is described earlier has employed for converting the issue of human action recognition into a typical HMM evaluation issue that is addressed with the help of standard algorithm

2.**Dynamic Bayesian Networks**: It is a Bayesian network having a similar structure that is unrolled within the time axis. The significant extension of the Dynamic Bayesian Network is that more than 1 random variable is comprised in its state space as compare to the Hidden Markov Model which consists of one single random variable. Therefore, the Hidden Markov Model is considered as a simplified Dynamic Bayesian Network having final random variables and fixed graph structures. The DBN is regarded as a hierarchical Bayesian network technique utilized to identify the interactions of an individual.

Discriminative Models: The types of Discriminative Models are given below:

1.**Support Vector Machines**: These are the classic classification algorithms of discriminative models and are extensively implemented in recognizing human action. The SVM is carried out to deal with the issue of separation of instances into two classes. This algorithm focuses on

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:45:36 UTC from IEEE Xplore. Restrictions apply.

investigating the hyperplane that is useful to enlarge the margin of both classes.

2.Deep Learning Architectures: Generally, the classification of DL algorithm is done into 4 groups such as DNNs, ConvNets, RNNs and some emergent models. The Convolutional Neural Network is gaining the extensive utilization amongst the earlier defined DL model. In the perception of modality of the input data, these models have potential for learning the human action attributes from enormous modal data. The earlier techniques are planned on the basis of image data and assisted in modelling the human actions from the appearance and motion information related to the temporal and spatial changes.

3.**Conditional Random Fields**: These are the undirected graphical algorithms utilized to characterize the conditional probability of a specific label sequence denoted by Y, given a sequence of observations revealed by X in a compact form. The Hidden Markov Models are compared with others in order to recognize the activity. It is observed in the comparative analysis that the CRF are trained discriminatively and proved more efficient in comparison with Hidden Markov Model though its attributes are available according to the independence assumptions of the Hidden Markov Model [6]. The essential difference is described amid Hidden Markov Models and Conditional Random Fields. The HMMs suppose that observations are independent given their labels. Therefore, the assumption of this model is invalidated due to the complex features and thereafter, HMM is not considered as a proper generative model. This assumption of Hidden Markov Models (HMM) is discarded in Conditional Random Fields. Hence, any independence assumptions are not necessitated among the observation variables.

B. *Action Detection*

Researchers made use of DT features for detecting actions in time dimensions on the basis of robust efficiency of DT (Dense Trajectories) in action classification. The DL methods have a great impact on some methods which are useful to detect the action in a video on the basis of DL [7]. These techniques aim at expanding the range of object classification for detecting the action using Region-Based Convolutional Neural Networks (R-CNN), three-dimensional convolutional networks, and RNNs. Following are the two popular DL-based action detection methods:

1.**3D convolutional networks**: This approach makes use of three 3D convolutional networks for action detection. Initially, the video is segmented in the time dimension using the candidate region extraction network. After this, actions are detected using the classification network and the location network on the basis of extracted candidate segmentation areas.

2.**Multimodality multitask recurrent neural network**: This approach performs modeling of action in colored and skeleton data and integrates the action classification process. The technique is a proposed system that is based on the 3 D skeleton for understanding the connection between human body parts. The time-varying trajectory is generated which can help to recognize the human gestures. The sparse coding method generates the atoms which can compare the labeled data and generate the final output.

C. Phases of research work

1. **Data set Collection**: - A group of thirty volunteers having ages between 19 to 48 years is considered to carry out the experimentation. Each person performed six activities (WALKING, WALKING UPSTAIRS, WALKING DOWNSTAIRS, SITTING, STANDING, LAYING)

2. **Training of System**: - The process of extracting a set of values called features from an image is called feature extraction. Various features such as color, texture, morphological, and color coherence vector are generally utilized to identify human activity. These methods include such as GLCM, color co-occurrence technique, SGLDM, and histogram-based feature extraction. The Grey Level Co-occurrence Matrix is utilized to classify the texture.

3. **Classification**: The classification technique is implemented in which an input is taken from the training and testing set. The whole dataset will be divided into 60 and 40 ratios. The 60 percent will be the training set and 40 percent will be the test set for the activity prediction. KNN classifier is an instance-based classification model.

4. **Performance Analysis**: In this phase, the various parameters like accuracy, precision, and recall will be calculated for the performance analysis shown in Figure3

III. EXPERIMENTAL RESULTS

After importing libraries and Preprocessing UCI HAR DATA SET used for cleaning data set in which basically, we have

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:45:36 UTC from IEEE Xplore. Restrictions apply.

two set one is attribute set that have two parameters location and time and another one is target set based on result of attribute. After Dividing dataset into training data that have target set as well as attribute set and testing set that have only attribute set the classification is applied that gives two input (training and testing data) and its output based on predictor set that predicts on behalf of target set with predicting the validation and training set result by making Confusion matrix, training error is 0.8 and testing error after applying same procedure on testing data set is 0. 1751.Transposing help to change the dimensionality of output in a very effective manner during aggregation of signals by a combination of sample/timestep since the resulting shape is 7352 train /2947 test samples,128 timesteps, 9 signals. Training error 0.18 and validation error 0. 0185.similar to the testing set results, the testing error 0.0719. hence, accuracy improved and results the visualization in a very accurate manner

```
 Dimension of Train set (7352, 561)
Dimension of Test set (2947, 561)

Number of numeric features: 561
0.9936049303086993
Best score for training data: 0.9863982589771492

Best C: 100

Best Kernel: rbf

Best Gamma: 0.001
```

Figure 3: performance analysis chart

	precision	recall	f1-score	support
LAYING	0.99	1.00	1.00	537
SITTING	0.95	0.88	0.91	491
STANDING	0.90	0.96	0.93	532
WALKING	0.96	0.98	0.97	496
WALKING_DOWNSTAIRS	0.98	0.91	0.95	420
WALKING_UPSTAIRS	0.93	0.96	0.94	471
accuracy			0.95	2947
macro avg	0.95	0.95	0.95	2947
weighted avg	0.95	0.95	0.95	2947

F1-SCORE: gives TEST accuracy, a Ratio parameter

Support: describe number of actual occurrences of the class in the specified dataset

Recall: fraction of relevant instances

Figure 4: Data set training vs testing set dimension representation With accuracy 98% of training data score

```
Training set score for SVM: 1.000000
Testing  set score for SVM: 0.951137
```

Figure 5: SVM score of prediction from dataset giving 95% score.

```
Number of numeric features: 561
0.9936049303086993
Best score for training data: 0.9863982589771492

Best C: 100

Best Kernel: rbf

Best Gamma: 0.001

[[537   0   0   0   0   0]
 [  3 433  54   0   0   1]
 [  0  23 509   0   0   0]
 [  0   0   0 486   4   6]
 [  0   0   0   7 384  29]
 [  0   0   0  15   2 454]]
```

Figure 6: CONFUSION MATRIX representing the performance of classifier

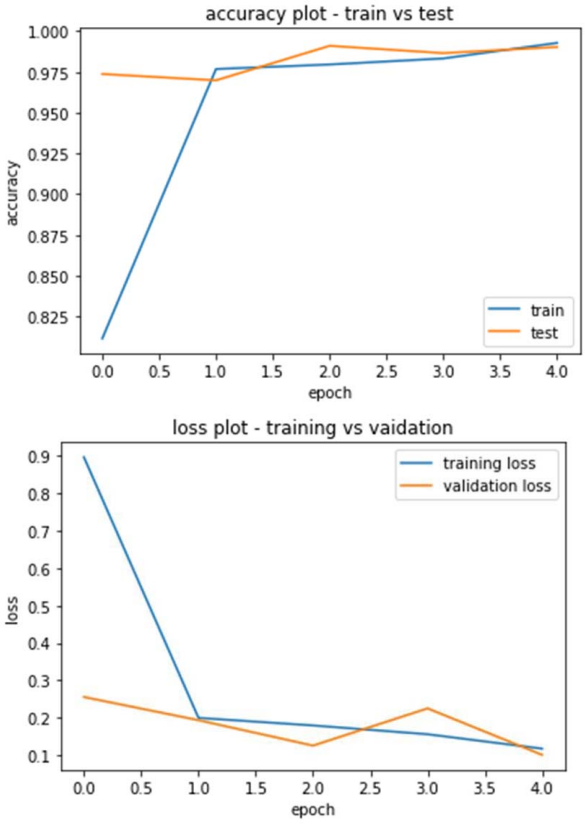


Figure 7: Accuracy and Loss Plot with respect to the epoch (set of time)

The epochs are a hyperparameter that defines the amount times that the algorithmic program can run through the whole training dataset.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:45:36 UTC from IEEE Xplore. Restrictions apply.

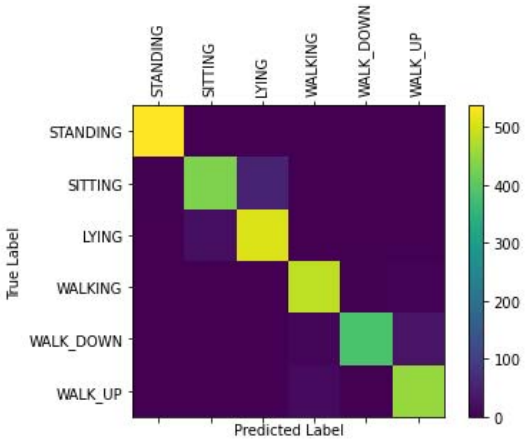


Figure 8: confusion matrix final model plot
(X axis: Predicted label, Y axis: true label)

Confusion matrix is based on two labels predictor label and true label that have four possibilities

1. True positive
2. True negative
3. False positive
4. False negative

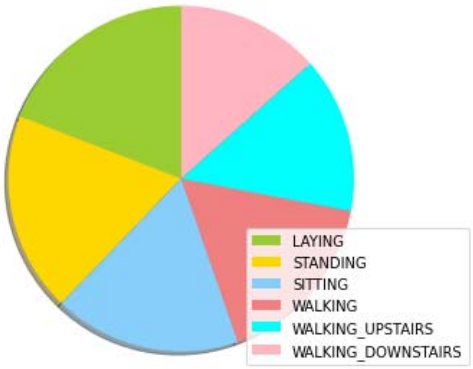
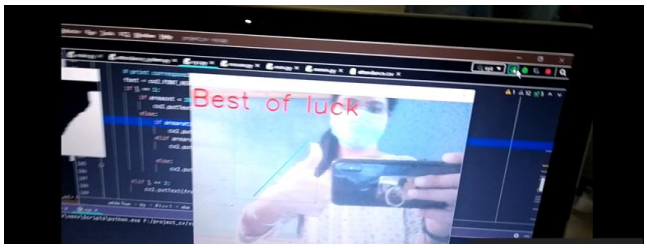


Figure 9: Classification report using Pi chat visualization



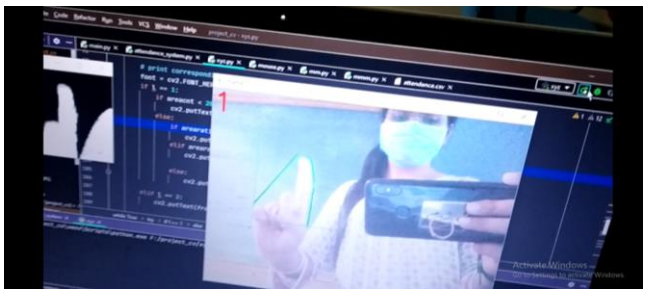




Figure 10**:** Capture of different gesture (BEST OF LUCK,1,2) through webcam showing accuracy with high precision

IV. CONCLUSION

This research work is based on human activity detection. To detect the human activity the feature of each activity needs to be extracted using a feature extraction algorithm that consider maximum impact of a feature on target set to find out that particular attribute. The textural and eigen features of the input image are extracted for human activity detection. The existed method rely completely based on deep learning architecture which is time taking process as it takes time on training that further increase testing time so there is need to decrease this time consumption and increasing the accuracy accordingly The proposed technique will improve performance for human activity detection in terms of many parameters like accuracy, precision, and recall. The whole dataset is divided into training and testing. The testing target set value will be compared with actual values which will improve accuracy for the human activity detection. Basically, designing a system that have less training time. The proposed technique is also compared with the other technique which authenticates the reliability of the model.

FUTURE WORK: The technique of 3-D Skelton is used for feature extraction. The hybrid classification method is designed in this research work which improves accuracy for human activity detection. Further work could be done using this hybrid technology in which more than 3 classifiers could be used to make it more accurate along with the use of different embedded sensor applications.

Limitations of Proposed work: Author did not find any limitation of proposed work.

ACKNOWLEDGEMENT

The first author would like to thank Delhi Technological University for providing lab facility and faculty for guidance.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:45:36 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] Nilam Nur Amir Sharif, Siti Mariyam Shamsuddin, "Human action invariance for human action recognition", 2015,9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)

[2] Earnest Paul Ijjina, C. Krishna Mohan, "Human Action Recognition Based on Recognition of Linear Patterns in Action Bank Features Using Convolutional Neural Networks", 2014, 13th International Conference on Machine Learning and Applications

[3] Jia Liu, Jie Yang, Yi Zhang, Xiangjian He, "Action Recognition by Multiple Features and Hyper-Sphere Multi-class SVM", 2010, 20th International Conference on Pattern Recognition

[4] Suraj Vantigodi, R. Venkatesh Babu, "Real-time human action recognition from motion capture data", 2013, Fourth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)

[5] Zi-Ming Wu, Wing W. Y. Ng, "Human action recognition using action bank and RBFNN trained by L-GEM", 2014, International Conference on Wavelet Analysis and Pattern Recognition

[6] Tongyang Liu, Yang Song, Yu Gu, Ao Li, "Human Action Recognition Based on Depth Images from Microsoft Kinect", 2013, Fourth Global Congress on Intelligent Systems

[7] Wei Huang, Q. M. Jonathan Wu, "Human action recognition based on Self Organizing Map", 2010, IEEE International Conference on Acoustics, Speech and Signal Processing

[8] Xinyi Liang, Hong-Bo Zhang, Yi-Xiang Zhang, Jin-Long Huang, "JTCR: Joint Trajectory Character Recognition for human action recognition", 2019, IEEE Eurasia Conference on IoT, Communication, and Engineering (ECICE)

[9] Qinkun Xiao, Yang Si, "Human action recognition using autoencoder", 2017, 3rd IEEE International Conference on Computer and Communications (ICCC)

[10] Thien Huynh-The, Cam-Hao Hua, Nguyen Anh Tu, Jae-Woo Kim, Seung-Hwan Kim, Dong-Seong Kim, "3D Action Recognition Exploiting Hierarchical Deep Feature Fusion Model", 2020, 14th International Conference on Ubiquitous Information Management and Communication (IMCOM)

[11] B Jagadeesh, Chandrashekar M Patil, "Video-based action detection and recognition human using optical flow and SVM classifier", 2016, IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)

[12] Amir Nadeem, Ahmad Jalal, Kibum Kim, "Human Actions Tracking and Recognition Based on Body Parts Detection via Artificial Neural Network", 2020, 3rd International Conference on Advancements in Computational Sciences (ICACS)

[13] Yun Han, Sheng-Luen Chung, Arul Murugan Ambikapathi, Jui-Shan Chan, Wei-You Lin, Shun-Feng Su, "Robust Human Action Recognition Using Global Spatial-Temporal Attention for Human Skeleton Data", 2018, International Joint Conference on Neural Networks (IJCNN)

[14] Sharmin Majumder, Nasser Kehtarnavaz, "Vision and Inertial Sensing Fusion for Human Action Recognition: A Review",2020, IEEE Sensors Journal

[15] Chen Chen, Roozbeh Jafari, Nasser Kehtarnavaz, "A Real-Time Human Action Recognition System Using Depth and Inertial Sensor Fusion", 2016, IEEE Sensors Journal

[16] Hang Yan, Beichen Hu, Gang Chen, E. Zhengyuan, "Real-Time Continuous Human Rehabilitation Action Recognition using OpenPose and FCN", 2020, 3rd International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)

[17] Naresh Kumar, Nagarajan Sukavanam, "Human Action Recognition from Motion Trajectory using Fourier Temporal Features of Skeleton Joints", 2018, International Conference on Advances in Computing and Communication Engineering (ICACCE)

[18] Suraj Prakash Sahoo, Samit Ari, Kamalakanta Mahapatra, Saraju P. Mohanty, "HAR-Depth: A Novel Framework for Human Action Recognition Using Sequential Learning and Depth Estimated History Images", 2020, IEEE Transactions on Emerging Topics in Computational Intelligence

[19] Carlos Roig, Manuel Sarmiento, David Varas, Issey Masuda, Juan Carlos Riveiro, ElisendaBou-Balust, "Multi-Modal Pyramid Feature Combination for Human Action Recognition", 2019, IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)

[20] Xueping Liu, Yibo Li, Your Li, Shi Yu, Can Tian, "The Study on Human Action Recognition with Depth Video for Intelligent Monitoring", 2019, Chinese Control and

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:45:36 UTC from IEEE Xplore. Restrictions apply.

Decision Conference (CCDC)

[21] Meng-Lin Chiang, Jian-Kai Feng, Wen-Lin Zeng, Chiung-Yao Fang, Sei-Wang Chen,"A Vision-Based Human Action Recognition System for Companion Robots and Human Interaction", 2018, IEEE 4th International Conference on Computer and Communications(ICCC)

[22] Tin ZarWint Cho, May Thu Win, Aung Win, "Human Action Recognition System based on Skeleton Data", 2018, IEEE International Conference on Agents (ICA)

[23] Peng Wang, Yuliang Yang, Wanchong Li, Linhao Zhang, Mengyuan Wang, Xiaobo Zhang, Mengyu Zhu, "Research on Human Action Recognition Based on Convolutional Neural Network", 2019, 28th Wireless and Optical Communications Conference (WOCC)

[24] Sui Longfei, Miao Lanfang, Li Zhifei, "Human Body Action Recognition Based on Bone Information by Kinect", 2018, International Computers, Signals, and Systems Conference (ICOMSSC)

[25] Sheng Li, Tingting Jiang, Yonghong Tian, Tiejun Huang, "3D Human Skeleton Data Compression for Action Recognition", 2019, IEEE Visual Communications and Image Processing (VCIP)

[26] Chengwu Liang, Deyin Liu, Lin Qi, Ling Guan, "Multi-Modal Human Action Recognition with Sub-Action Exploiting and Class-Privacy Preserved Collaborative Representation Learning", 2020, IEEE Access

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:45:36 UTC from IEEE Xplore. Restrictions apply.

Semiconductor Science and Technology

ACCEPTED MANUSCRIPT

A 32 nm single ended single port 7T SRAM for low power utilization

To cite this article before publication: Bhawna Rawat *et al* 2021 *Semicond. Sci. Technol.* in press https://doi.org/10.1088/1361-6641/ac07c8

Manuscript version: Accepted Manuscript

Accepted Manuscript is "the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an 'Accepted Manuscript' watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors"

This Accepted Manuscript is **© 2021 IOP Publishing Ltd**.

During the embargo period (the 12 month period from the publication of the Version of Record of this article), the Accepted Manuscript is fully protected by copyright and cannot be reused or reposted elsewhere.
As the Version of Record of this article is going to be / has been published on a subscription basis, this Accepted Manuscript is available for reuse under a CC BY-NC-ND 3.0 licence after the 12 month embargo period.

After the embargo period, everyone is permitted to use copy and redistribute this article for non-commercial purposes only, provided that they adhere to all the terms of the licence https://creativecommons.org/licences/by-nc-nd/3.0

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions will likely be required. All third party content is fully copyright protected, unless specifically stated otherwise in the figure caption in the Version of Record.

View the article online for updates and enhancements.

This content was downloaded from IP address 14.139.251.100 on 11/06/2021 at 06:39

A 32 nm single ended single port 7T SRAM for low power utilization

Bhawna Rawat[1] and Poornima Mittal[1*]

[1] Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India

E-mail: bhawnarawat12@gmail.com, *poornimamittal@dtu.ac.in (Corresponding Author)

Received xxxxxx
Accepted for publication xxxxxx
Published xxxxxx

Abstract

In this paper a seven-transistor SRAM bit-cell with a single bitline architecture is proposed. This cell is designed at 32 nm and is operational at 300 mV. The static-noise-margin for read and hold mode is 90 mV, while write margin is 180 mV. Monte Carlo analysis for 6σ global variations and temperature variation analysis for temperature ranging from -10 ºC to 80 ºC, validates its performance. The cell is compared against other single ended 5T, 6T, 7T, 8T, 9T, 10T SRAM cells and is found to be superior in performance to other bit-cells. As the leakage current is low, the I_{ON}/I_{OFF} ratio is high in comparison to its other counter-parts. The power consumption of the bit cell is also found to minimal for all modes of operation. While the dynamic write analysis demonstrates that the proposed cell completes write operation in 10 ns pulse width. Moreover, the improvement in the performance is obtained for area as low as 0.539 μm^2. The area of 5T, 6T, 7T-1, 7T-2, 7T-4, 7T-5, 8T, 9T and 10T cell is greater than 7TP bit cell area by 22.17, 51.8, 35.8, 13.9, 30.4, 6.78, 56.6, 63.3 and 75.5 %.

Keywords: Leakage Power, low power, I_{ON}/I_{OFF} ratio, single ended, single port, 7T SRAM

1. Introduction

Portable devices with greater processing capabilities are vogue. Digital designers are designing System on Chip (SoC) that will satisfy the ever increasing processing thirst of the user. SoCs integrated with internet of things (IoT) setups or those used for artificial intelligence applications are required to be highly reliable. The core of these SoC's are formed by processors. The main attribute of a processor is its multi-functionality and high speed. For faster application, each processer is dependent on an internal cache. This cache is a high speed memory dedicatedly available for the functioning of the processor. Cache is an array of static random access memory (SRAM) bit cells. Therefore, optimizing the functioning, area and power requirement of the bit cell has become the center of attention for memory designers.

Another aspect that makes optimization of SRAM bit cells important is its large area footprint on a chip. Almost 30 % of the total power consumption of an embedded system is due to SRAM [1]. While, ITRS predicts that in the near future 90% area of a processor will be occupied by cache [2]. With the present-day advancement trends, it is optimal to design devices that have longer battery life. Therefore, reduction is power consumption of SRAM bit cells is of utmost importance. For SRAM bit cell, the power consumption can be bifurcated into static and dynamic power consumption. Dynamic power in SRAM is caused by the charging/discharging of bitlines, which accounts for 60% of the total dynamic power consumption of cache [3]. For a differential bit cell one bitline always discharges during read and write operations. Therefore, to reduce leakage and switching power consumption of the bit cell, single ended SRAM bit cells are gaining attention [4-7]. Scaling the supply voltage (V_{DD}) is another convenient method to reduce dynamic power consumption [8]. Many other configuration for SRAM bit cells have been reported to rectify the abovementioned problem [9-15]. But reduction in V_{DD} is limited by process-voltage-temperature (PVT) variations in short channel devices [16]. Also, V_{DD} scaling degrades the static noise margin (SNM) because of its exponential relation with the threshold voltage (V_{TH}) [17]. The situation is further exacerbated with the increasing read write conflict inherent to the conventional cells. Adding to the pitfalls is the half select issues that leads to an unintentional flipping in the unselected SRAM bit cell in a selected array. To rectify the same various researchers have reported different cells that focus on improving (i) SRAM assist techniques to enhance the

Accepted Manuscript

performance of the bit cell during the read or the write operations, [18-22], (ii) isolating the read and write port to enable low voltage operation and also to eliminate the read-write conflict inherent in a SRAM bit cell [23-25] and (iii) separating the read path current from the memory storage node has also proved to improve the read stability for the bit cell [24-28]. Though these techniques have improved the performance, they have resulted in area and power penalties for the bit cells. Also, with the decreasing feature size, the reliability of the bit cell is threatened by bit error and soft error. An effective methodology to resolve these errors is to utilize bit interleaving (BI) technique. But, this is only applicable when the designed SRAM bit cell employs cross point cell selection method. This implies that the access transistors for an operation are controlled by different row and column based signals [29]. This configuration for the bit cell that utilizes the cross point cell selection method makes the bit cell free from half select issues. But, in doing so most cells tend to increase the transistor count resulting is high area footprint. Thereby reducing the economic feasibility for an SRAM bit cell.

In this work, a single ended, seven transistor SRAM bit cell is proposed. The proposed 7T (7TP) bit cell is dependent on a single bitline for its read and write operation. Thereby, decreasing the leakage and dynamic power consumption of the bit cell. Also, the cell uses independent row and column signals to control the read and write operation, to make the cell free from half select issue. Additionally, the read port of the cell is designed to not include the memory storage node into the read current path. Therefore, making the read operation for the bit cell highly stable. The highlight of the design is that it has a balanced hold, read and write margin. The rest of the paper is organized as follows – the previously reported SRAM bit cell are explained in section II. In section III the schematic and the operations for the proposed SRAM bit cell are elaborated. The static performance of the bit cell is justified in section IV. Section V and VI present variation tolerance analysis and the failure probability analysis of the cells. The I_{ON}/I_{OFF} ratio for all the bit cells is explained in section VII. The dynamic performance of the bit cell is expressed in section VIII. In section IX, the power performance of the cell is presented. Whereas, section X compares the layout and area footprint of the bit cells. A comprehensive conclusion to the work is presented in section XI.

2. Overview of the pre-existing SRAM bit cells

Several architectures and circuit designs for SRAM bit cells have been proposed to overcome diverse challenges, that an SRAM bit cell encounters. The design objectives such as – high bit cell density, low leakage area, faster operation, and low V_{DD} requirement – form the core of an SRAM bit cell design. Some bit cells that achieve the aforementioned objectives are presented in Fig. 1. The common ground for the bit cells discussed in this section is that they have a single ended approach for implementing their operation. The charging and discharging of the bit line is a major contributor to dynamic power consumption. Therefore, limiting one bitline for an operation reduces the power consumption for a memory cell significantly.

The simplest possible single ended SRAM bit cell consists of five transistors (5T) and was first proposed by Carlson *et al.* [30] in 2004. The circuit diagram for the bit cell is represented in Fig. 1 (a). The cell is composed of a conventional memory core; made up of two mutually coupled inverters (P1-N1 and P2-N2). The cell is dependent on a single access transistor; N3 for its read and write operations. A single ended bit cell is dependent on the consecutive effect of the inverters to perform the write operation. But, the inability of an NMOS based access transistor to pass a good '1', limits the write margin for write X = 1. This limitation is further exacerbated with technology scaling and reduction in V_{DD}.

The most popular and industry preferred SRAM bit cell is the conventional 6T (6TC) SRAM bit cell. But with reduction in technology node and V_{DD} scaling, this cell is unable to keep up optimal performance. Additionally, since 6TC bit cell is differential in nature, the power consumption is considerably high. Therefore, to reduce the charging/discharging events for an SRAM bit cell, a single ended 6T SRAM bit cell was reported by Surana and Mikie [31] in 2019. The circuit diagram for this single ended 6T (6T) SRAM bit cell is given in Fig. 1 (b). The memory core for the bit cell is composed of an inverter (P2-N2) and pull up device; P1. The pull down device; N1 has its gate and source shorted. Thereby, enabling N1 to always remain in cut-off region. This is done to improve the write margin for the bit cell by reducing the tussle between N3 and N1. But in doing so, the X = '0' margin is disturbed. To restore the same, a low threshold voltage (LV_{TH}) NMOS is used for N1. The memory core for the bit cell is accessed using two NMOS transistors; N3 and N4. Since the bit cell is single ended in nature, N3 and N4 are used for read and write operations respectively. Therefore, a single bit line is used effectively for an operation. This previous bit cell has reasonable performance but the area footprint and power consumption of the bit cell is high, owing to the large dimensions for the devices.

A seven transistor single ended SRAM bit cell, 7T-1, was presented by Aly and Bayoumi [32] in 2007. The circuit diagram for the bit cell is given in Fig. 1 (c). The bit cell is a slight modification of the conventional 6T SRAM bit cell. The memory element of the bit cell is a combination of two inverter pairs (P1-N1 and P2-N2) connected by a feedback connection. The feedback connection is enabled by an additional transistor; N5. The utility of N5 is to establish/break the feedback connection between the inverter pair. The discharging for the write operation cannot be eliminated, as it is necessary to flip the data of the bit cell. And in a conventional cell both the bitlines discharge making the bit cell highly vulnerable to noise and erroneous event. Using a single bit line for the write operation yields better results. Therefore, this bit cell follows a single ended write and a differential ended read operation. But this bit cell was designed for 180 nm technology node and 1.8 V V_{DD}. Therefore, at scaled technology node and V_{DD}, this bit cells suffers in performance. The SNM for read operation suffers the most due to its differential nature.

Accepted Manuscript

Therefore, a different 7T SRAM bit cell; 7T-2, was proposed by Yang *et al.* [33] in 2016. This bit cell uses a single ended approach for both read and write operations. The circuit diagram for this 7T-2 SRAM bit cell is given in Fig. 1 (d).This bit cell is dependent on transmission gates; N3 and P3, for disconnection of node X from input node for the right-side inverter. The cell is accessed through transistor; N4. Four control signals; WL, WWLA, WWLB and BL, are used to steer the bit cell into different modes of operations. But this bit cell is dependent on two pulses to perform the write operation. Since it is a single ended bit cell, this reduces the speed of write operation. Also, the use of additional control signals increases its complexity. This results in additional peripheral circuit, thereby increasing the overall footprint as well as power consumption of the cache memory.

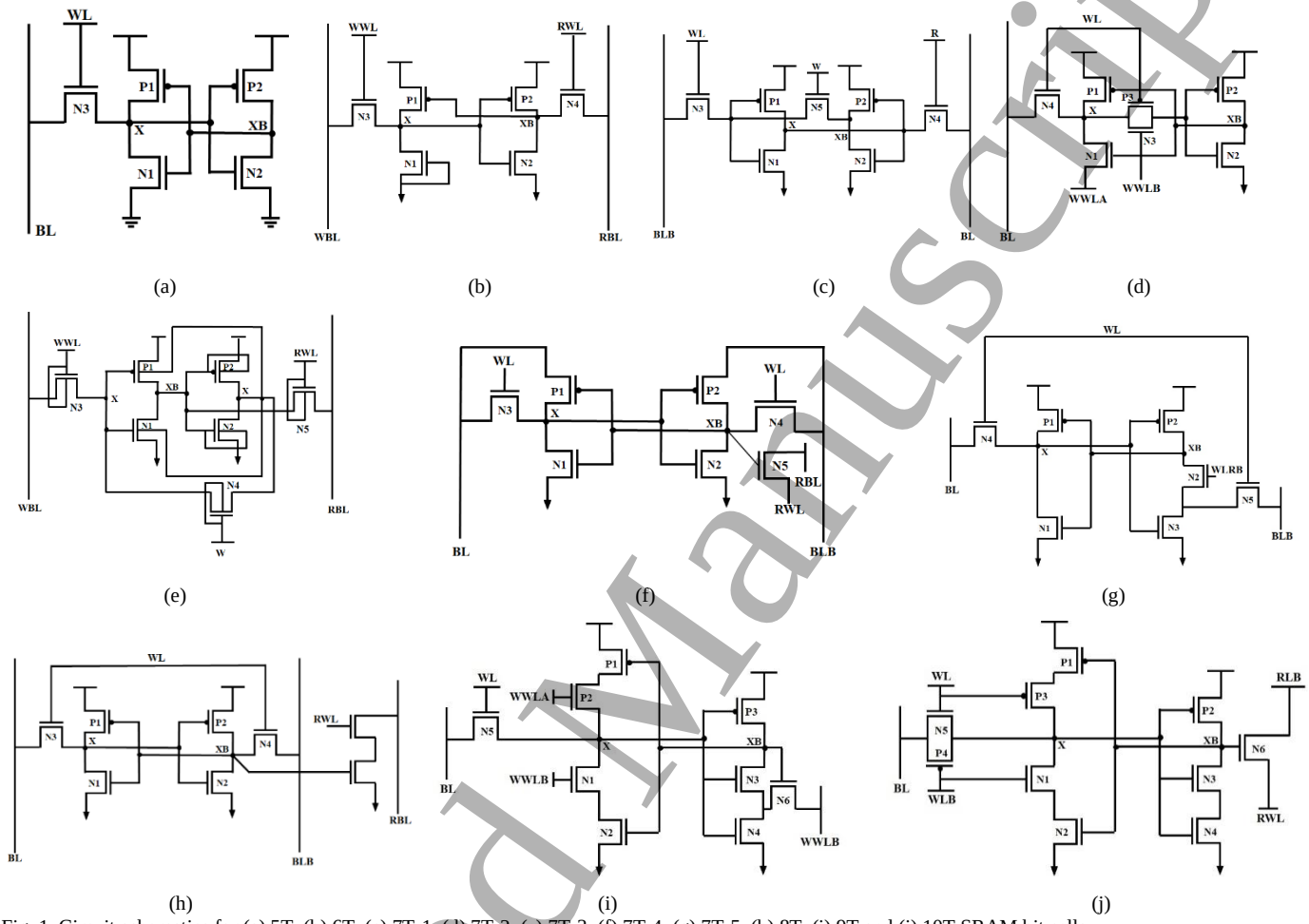


Fig. 1. Circuit schematics for (a) 5T, (b) 6T, (c) 7T-1, (d) 7T-2, (e) 7T-3, (f) 7T-4, (g) 7T-5, (h) 8T, (i) 9T and (j) 10T SRAM bit cells

In 2018, Sanapala *et al.* [34] designed a 7T single ended dual port SRAM bit cell (7T-3) utilizing a Schmitt Trigger circuit. The schematic for the bit cell is shown in Fig. 1 (e). This bit cell relies on dynamic body bias technique which allows changing the output switching voltage by dynamically varying the threshold voltage of the transistors. The highlight of this cell is the second inverter that acts as a body biasing circuit for the first inverter. This technique enables fast charging and discharging of the bit cell. Also, to enable better write ability, the cell is reliant on breaking the feedback connection between the inverter pair with the help of transistor; N4 and control signal; W. A flaw with this design is that node X has lesser strength compared to XB.

Another recent schematic for 7T SRAM bit cell was proposed by Ahmed *et al.* [24] in 2018. This cell is an improvement of the cell presented by Liu *et al.* [35]. This cell has an isolated read port structure. The schematic for the bit cell is depicted in Fig. 1 (f). The read operation for this cell is performed using an additional port formed by a single transistor; R2. For entire duration of the read operation, the access transistors remain OFF. Thus, the read operation does not upset the value stored in the memory core. This ensures the read stability of the bit cell. But, to rectify the write fail during low V_{DD}, Ahmed *et al.* [24] modified the V_{DD} assist. The pull up devices; P1 and P2 for the core memory element are powered by bitline; BL and complementary bitline; BLB respectively. This method provides a data dependent supply collapse during the write operation.

Yet another 7T SRAM bit cell was reported by Oh *et al.* [36] in 2021. This cell has a differential read and write operation. In comparison to the conventional 6T SRAM bit cell, the reported 7T uses an additional transistor; N3 controlled by the wordline right bar (WLRB) signal. The schematic for 7T-5 SRAM bit cell is presented in Fig. 1 (g). The N3 transistor forms a stacked inverter configuration for the bit cell, thereby increasing the power consumption for the

Accepted Manuscript

cell. But, this configuration increases the power consumption and reduces the ON current for the bit cell. To counterbalance the loss of differential read, the cell turns OFF the N3 transistor to improve the read static noise margin. The write operation for the bit cell is very strictly time constrained. If the WLRB signal is not carefully modulated with respect to WL signal, the cell may register write difficulty. This may also lead to an increase in write delay for the circuit.

Further, Chang *et al*. [37] proposed an 8T cell with single ended read operation to isolate the read and write operations, while maintaining the cross coupled nature of the memory cell. This, 8T SRAM bit cell has an additional read port to isolate the read and write operations. The schematic for 8T SRAM bit cell is illustrated in Fig. 1 (h). Here, the read port is made of two NMOS transistors; N5 and N6 stacked one over the other. The configuration of stacked read access transistor is used to reduce the bit line leakage [38]. The stacked read access transistors are accessed through read word line; RWL, to enable the read operation and read bit line; RBL, provides the discharge path. The gate of the lower NMOS transistor; N6, is connected to the XB node; whereas, the other transistor; R1, is controlled by RWL. The source of N5 is connected to the RBL. This 8T cell rectifies the read vulnerability and at the same time reduces the leakage of the 6TC SRAM bit cell. But, this is achieved at the price of high power consumption and increased silicon area.

Yet another single ended SRAM bit cell was reported by Cho *et al*. [39] in 2020. This cell has a 9T configuration and utilizes a Schmitt Trigger inverter configuration to boost the performance of the bit cell. The schematic for the 9T SRAM bit cell is given in Fig. 1 (i). The memory core of the bit cell is formed by stacked transistors (P1, P2, N1, and N2) and a Schmitt trigger-based inverter (P3, N3, N4, and N6). The bit cell is accessed through the NMOS transistor; N5. The N5 transistor is controlled by the wordline; WL signal. Whereas, the write wordline A; WWLA and write wordline B; WWLB (column-based signals) are used to control the write operation for the bit cell. The data to be written into the cell is placed on the bitline, BL. This cell has a stable static performance but its performance degrades with process variability also the area overhead for the bit cell is extremely high.

Eslami *et al*. [40] reported a cell similar to the aforementioned 9T SRAM bit cell in 2020. This cell has a 10T configuration and the schematic for the same is presented in Fig. 1 (j). The memory core for the bit cell is formed by stacked inverters. The left inverter is formed by P1, P2, N1, and N2, while the other inverter is formed by P3, N3, and N4. The write port for the bit cell is formed by N5 and P4 controlled by signal wordline (WL) and wordline bar (WLB) respectively. For the read operation, the cell is dependent on an isolated read port formed by NMOS transistor; N6. This bit cell improves the static stability for the SRAM bit cell in the nanometer range but suffers in terms of dynamic ability and the area overhead is extremely high.

In all these aforementioned SRAM bit cells, the basic memory element is a cross-coupled inverter pair. Other additional transistors are used either to decouple the read-write operation or to break the cross coupled feedback connection. Nevertheless, these bit cells have not taken any measure to incorporate process variation tolerance for improving the stability of the memory element for the bit cell. Additionally, as V_{DD} is being scaled, improving the stability of the inverter pair is becoming pivotal [41].

3. Circuit Design and Operation

A single ended, seven transistor SRAM bit cell is proposed. The circuit diagram for the proposed 7T (7TP) SRAM bit cell is given in Fig. 2. The 7TP bit cell is composed of a memory core and two access transistors; N3 and N4, to perform the write and read operations respectively. The memory core of the proposed bit cell is composed of two inverter pairs; inv1 (P1-N1) and inv2 (P2-N2). The output node of inv1; X, is connected to the input node of inv2.

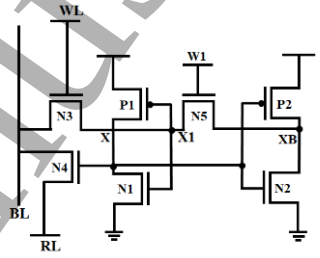


Fig. 2 Schematic diagram for the proposed 7T SRAM bit cell

The output node of inv2; XB in turn is connected to the input node of inv1, via an NMOS transistor; N5. The utility of N5 is to establish and eliminate the connection between XB and X1. The N5 transistor is controlled by an additional control signal; W1. This signal is low only during the write operation; while for read and hold operations, W1 is maintained high.

3.1 Read Operation

The proposed cell is dependent on N4 transistor to perform the read operation. The gate terminal of N4 is connected to the X node (output of inv1) for enabling the read operation. If the data stored at X is '1' then, during the read operation N4 is ON and BL discharges via N4 to RL. While for X = '0', no discharge is observed as the N4 transistor is OFF. During the read operation, the memory core of the bit cell is undisturbed.

As, the read discharge path for the bit cell does not include the memory storage node of the bit cell. Therefore, during read operation, the read discharge current does not impact the data stored in the memory core. An SRAM bit cell with such a read configuration has its RSNM equivalent to HSNM [27]. This read port configuration is also referred to as isolated read with read free noise margin [42].

An isolated read configuration for the SRAM bit cell implies that the write and read ports are independent of each other. The read operation for an SRAM bit cell is dependent on discharging of the bitline to be detected by the sense

Accepted Manuscript

amplifier. Some bit cells, such as reported by Kushwah and Vishvakarma [43], Pasandi and Fakhraie [44] and Surana and Mikie [31] though have a separate port for read operation, but have the read discharge current path though the memory storage node. Therefore, the discharging process during the read mechanism has the potential to disturb the data stored at the memory storage node. But other cells, such as reported by Ahmed *et al.* [28], Wen *et al.* [45], and Sil *et al.* [26] isolate the memory storage node from the read path. For these cells the read current path does not pass through the memory storage node. Thereby, these memory cells are not disturbing the data already stored in the bit cell.

Thus, the HSNM and RSNM for the proposed cell are found to be equal. Also, the read operation for the bit cell is controlled by independent row and column based signals, BL and RL respectively. Thus making the cell free from half select issues. Therefore, it can be inferred that the RSNM free isolated read configuration for the 7TP SRAM bit cell also makes it immune to half select issues [29]. Also, BL for the read operation will only discharge when the data stored in the bit cell is '1'. Thereby reducing the discharging event for the bit cell.

For a conventional SRAM bit cell, half-select disturbance occurs when an unintentional data flip is registered for an unselected bit cell. This unselected SRAM bit cell shares a common bitline or wordline with a selected bit cell (i.e. the unselected and selected cells lie in the same row or column) [28]. Ahmad *et al.* [28] in 2018 reported two 11T SRAM bit cells that are completely half select disturbance free. Alike these two cells, the 7TP SRAM bit cell has isolated read SNM free port. Also, the read port for the 11T SRAM bit cell reported by Ahmad *et al.* [28] and our 7TP SRAM bit cell have two independent row and column based signals to perform the read operation. In the 7TP SRAM bit cell, the RL and RWL signals make the 7TP SRAM bit cell half select free during the read operation. Similar, propositions are also made in by Shin *et al.* [46] in 2017. This suggests that the most efficient way to achieve half select free operation is to use cross point cell selection, where the operation is controlled by different row and column based signals [29, 44]. The 7TP SRAM bit cell employs the same technique of row and column signals to access the bit cell during the read as well as write operation.

3.2 Write Operation

The write operation for the bit cell is enabled by turning ON the N3 transistor by setting WL high. While the N5 transistor is turned OFF by setting the W1 signal low. Complement of the data to be placed into the memory core is placed on BL (i.e if '0' is to be written into the memory, the BL is set at '1'). The data then gets transferred from BL to X1 node (input of inv1) via N3. X1 node then drives inv1 to develop the desired value at node X. Node X in turn drives inv2 to generate the value at XB. As the bit cell has a single ended write operation, it is dependent on the mutual effect of cascaded inverter pair to perform the write operation. Therefore, to enhance the write ability for the cell, a LV_{TH} NMOS; N1 is used to form the pull-down network for inv1. LV_{TH} MOS ensures faster write operation for inv1, thereby resulting in an early availability of data at input on inv2.

During the write operation, the memory core of the bit cell is accessed using the transistor; N3, controlled by signal; WL. The other transistor that controls the write operation for the bit cell is the N5 transistor, controlled by the W1 signal. These two signals form row and column based cross point cell selection signal during the write operation. Thereby making the cell half select free during the write operation. Thus utilizing independent row and column based signals to control the read and write operation help in eliminating the half-select disturbance for the SRAM bit cell [27].

3.3 Hold Operation

During the hold mode, WL is held low, while RL and W1 signals are maintained high. This enables the N3 transistor to OFF state, while the N5 transistor is turned ON. Thus the effective circuit during the hold operation for the proposed 7T SRAM bit cell is that of a cross coupled inverter pair, disconnected from the BL.

4. Static Analysis

All the SRAM bit cells discussed in section II along with the 7TP SRAM bit cell are simulated and subsequently, the results for static analysis are explained in this section. The circuits of all SRAM bit cell are designed at 32 nm technology node using PTM library files. For the entire duration of the study, the temperature for the simulation environment is maintained at 27 ºC.

A highly reliable SRAM bit cell has a high SNM for the bit cell. So, all the bit cells are foremost compared in term of their static margins. The read and hold SNM for the bit cell are measured using the butterfly curve, obtained by superimposing the voltage transfer characteristics for the bit cells during the hold and read mode respectively. The side of the square inside the lobe of the butterfly curve is taken as the measure for SNM [42]. The butterfly curve for the read SNM (RSNM) and hold SNM (HSNM) for the proposed SRAM bit cell are overlapping and are given in Fig. 3(a).

A read SNM free isolated read does not discharge through the memory core of the bit cell, as is observed in a conventional SRAM bit cell. For the proposed 7T SRAM bit cell also, the read current path for the bit cell does not include X or XB node, thereby enabling the cell to be read-SNM free [26, 41]. The read port for the bit cells decouples the storage node (X and XB) from the bitline using the read access transistor. This produces a high RSNM and nearly equivalent to the HSNM [27, 10]. Ahmad *et al.* [28] in 2017 have reflected the same in their paper as the HSNM and RSNM are not mentioned independently but categorized as SNM for the bit cell. This is done as the read current path for the proposed 11T bit cells is exclusive of the memory storage node. The same is also reported by Sil *et al.* [26] in 2008. Additionally, He *et al.* [27] in 2019 also graphically represented the same in their work. Thus, the HSNM and

Accepted Manuscript

RSNM for the proposed cell are found to be equal. The proposed 7T SRAM bit cell has a single ended configuration, wherein, the read path is not passing through the memory core for the bit cell. Therefore, the HSNM and RSNM for the bit cell are also equivalent. The comparison for SNM of the 7TP SRAM bit cell with other bit cells is shown in Fig. 4.

The stability of the write operation is measured using the parameter – write margin (WM). Traditionally, write stability was measured using the write SNM. But recent studies has shown that WM is a more reliable parameter to ensure write stability [47]. WM is calculated for each SRAM bit cell using the method explained in [42]. The curve obtained for calculation of WM for the 7TP SRAM bit cell is given in Fig. 3(b). The values corresponding to the WM obtained for all the SRAM bit cells are graphically compared in Fig. 4. For the ease of comparison the values for HSNM, RSNM and WM for each cell are tabulated in Table I.

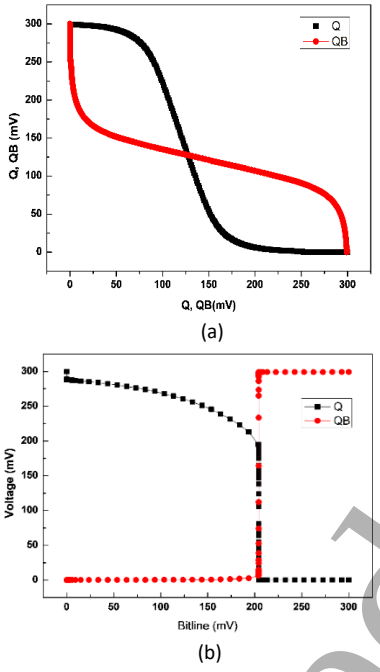


Fig. 3. The butterfly curve for (a) HSNM and RSNM, (b) WM for the proposed 7T SRAM bit cell

The 5T and 6T cells have a single ended read and have their discharge path pass through the memory storage node, thereby, leading to a collapse of butterfly curve for RSNM values (represnted by butterfly curve collapse (BCC) in table). Thus, suggesting a read failure event for the two SRAM bit cells. For the 5T cell, the WM is also considerably low at 50 mV, whereas the WM for 6T is considerable at 106 mV. The 7T-1 cell has a highly skewed performance, the HSNM is good but the RSNM is poor and the WM is extremely high at 75, 30 and 196 mV respectively. The 7T-2 cell though has a comparable HSNM and RSNM, but its write operation is a two-cycle process (represneted by Dual Pulse Write (DP) in the table) and therefore, cannot be compared with the other SRAM bit cell, as the other cells are able to perform the write operation in a single cycle. Two 7T SRAM configurations 7T-3 and 7T-4 though have comparable HSNM and RSNM performance with 7TP, but fail during the WM analysis at 300 mV. While, for 7T-5 SRAM bit cell the HSNM value is quite high, but its RSNM and WM metric are inferior at 18 and 105 mV.

TABLE I – Static Analysis Results

SRAM bit cell	HSNM (mV)	RSNM (mV)	WM (mV)
5T	80	BCC	50
6T	51	BCC	106
7T-1	75	30	196
7T-2	80	80	DP
7T-3	32	32	DP
7T-4	81	81	DP
7T-5	84	18	105
8T	84	84	144
9T	67	67	160
10T	83	83	182
7TP	90	90	180

BCC – Butterfly Curve Collapsed, DP – Dual Pulse Write

The 8T, 9T and 10T SRAM bit cell have a comparable static performance with the 7TP SRAM bit cell, but there area panelties are high. Also, the power consumption for the three cells is also high, because the 8T cell has a dual port configuration, and 9T and 10T have a stacked inverter configurations. The stacked inverter configuration also reduces the I_{ON}/I_{OFF} ratio for the two cell. Thereby, decreasing the number cells that can be placed in an array. Therefore, it can be inferred that 7TP SRAM bit cell has better static performance in comparison to the other bit cells.

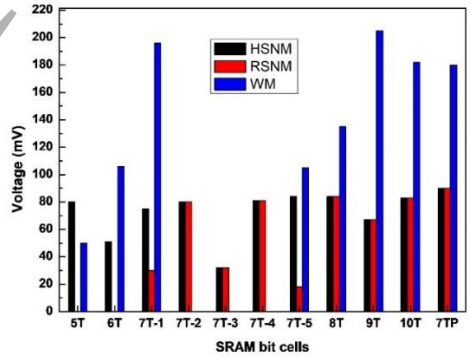


Fig. 4 Graphical representation for HSNM and RSNM for all SRAM bit cells

5. Tolerance Variation Analysis

The recent decade has observed a very high scaling trends in technology node. As most bit cells in this study were also proposed for higher technology node and now are being evaluated for scaled technology node. The foremost impact of technology node scaling is that it introduces serious variability issues in the circuit performance [45]. The perfromacne of SRAM bit cells is severely limited by PVT variations in deep sub micrometer regions [48,49]. Also, with the decreasing technology node the probability of error due to alpha particles and cosmic radiations has increased [50]. These errors are referred to as soft errors. As the 7TP SRAM bit cell is free from half select disturbance, BI with error correcting code is an effective method for dealing with

Accepted Manuscript

such errors [51]. A cells susceptibility to such errors further increases with process and temperature variation.

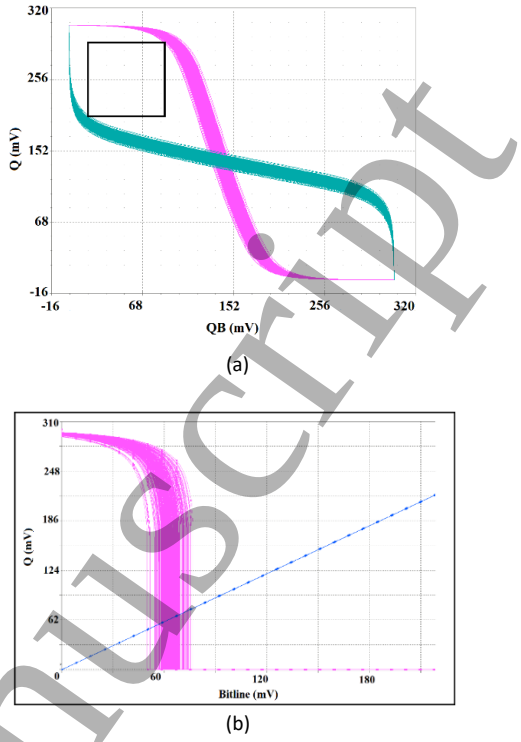
5.1 Process Variation Analysis

The increasing demand for larger SRAM based cache memory has pushed the designers towards lower technology nodes. So, designing nano-meter sized SRAM bit cell requires validation for performance in terms of variability issues.

With the increase in stress time, the bias temperature instability (BTI) increases the V_{TH} of the transistor and consequently degrades the performance of the circuit [28]. The BTI induced degradation of SRAM bit cell is different from NMOS and PMOS. In NMOS transistor, the BTI effect is caused by negative bias (NBTI) and for PMOS transistor the same is induced by positive bias (PBTI). But, both the effects impact the V_{TH} for the devices [52]. Therefore, researchers have modelled the impact of stress and aging on transistor in terms of variation in the V_{TH}. The shifts in V_{TH} of the transistor leads to uncertainty in the circuit behavior [53]. Jin *et al.* [52] and Shah *et al.* [53] have mentioned different equations for BTI variation but they model it in terms of V_{TH}.

Other factors that limit the performance of the bit cell are; hot carrier injection, gate oxide breakdown, backend dielectric breakdown, electro migration, and stress-induced voiding. But these factors have a random behavior and do not have a defined model of analysis. All variations can be transalted to an effective variation in the V_{TH} [54, 55]. Consequently, to analysie the impact of variation on device, the Gausian distribution for V_{TH} is an effective approach [56]. Therefore, for Monte Carlo simulation, the variation in the V_{TH} is taken as the primary factor for analysis. Ahmad *et al.* [28] in 2018 have used Monte Carlo simulations with a sample size of 5000 to analyze the impact of BTI on the bit cell. Hence, it is essential to explore variation tolerant design solutions to mitigate variability issues due to temporal degradations. The 7TP SRAM bit cell is analyzed for variation tolerance under the process variation analysis using Monte Carlo simulations for 10,000 sample size.

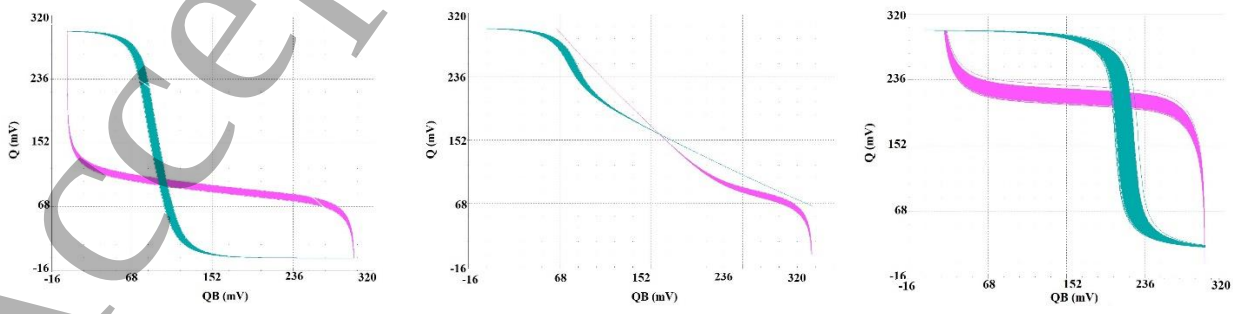
Statistical methods were employed to identify the range of variability in the performance of the 7TP SRAM bit cell. The main parameter for the statistical analysis is the threshold, V_{TH} mismatch. To assure that the yield loss would be less than 0.1 for a 10 Mb cache memory, the permissible failure limit, P_{fail}, should be less than 10^{-10} [57].



(a) (b)

Fig. 5. Monte Carlo Simulations for (a) HSNM and RSNM, (b) WM for the proposed SRAM bit cell

The nature of the probability curve is Gaussian so it corresponds to a μ-6σ factor. Here μ is the mean value of V_{TH}, while the 6σ factor represents the standard deviation of the distribution. So, the proposed SRAM bit cell is analyzed for 6σ variations for 10000 point Monte Carlo simulations. Therefore, Monte Carlo simulations are used and V_{TH} for the bit cell is varied around the mean value for 6σ variations. The HSNM and RSNM for 10,000 points Monte Carlo simulations of the 7TP SRAM bit cell are presented in Fig. 5. The outcomes of the process variation analysis reveal that the HSNM and RSNM for the 7TP SRAM bit cell can vary up to 75 mV for the worst-case situation. While the WM for the bit cell increases to 215 mV. This variation in the bit cell performance is for maximum deviation from the mean value. The variation in HSNM and RSNM is 16.6% while, WM records a maximum variation of 19.44% for the 7TP SRAM bit cell. The Monte Carlo simulation results for butterfly curve of all the cell described in section 2 are presented in Fig. 6.



(a) (b) (c)

Accepted Manuscript

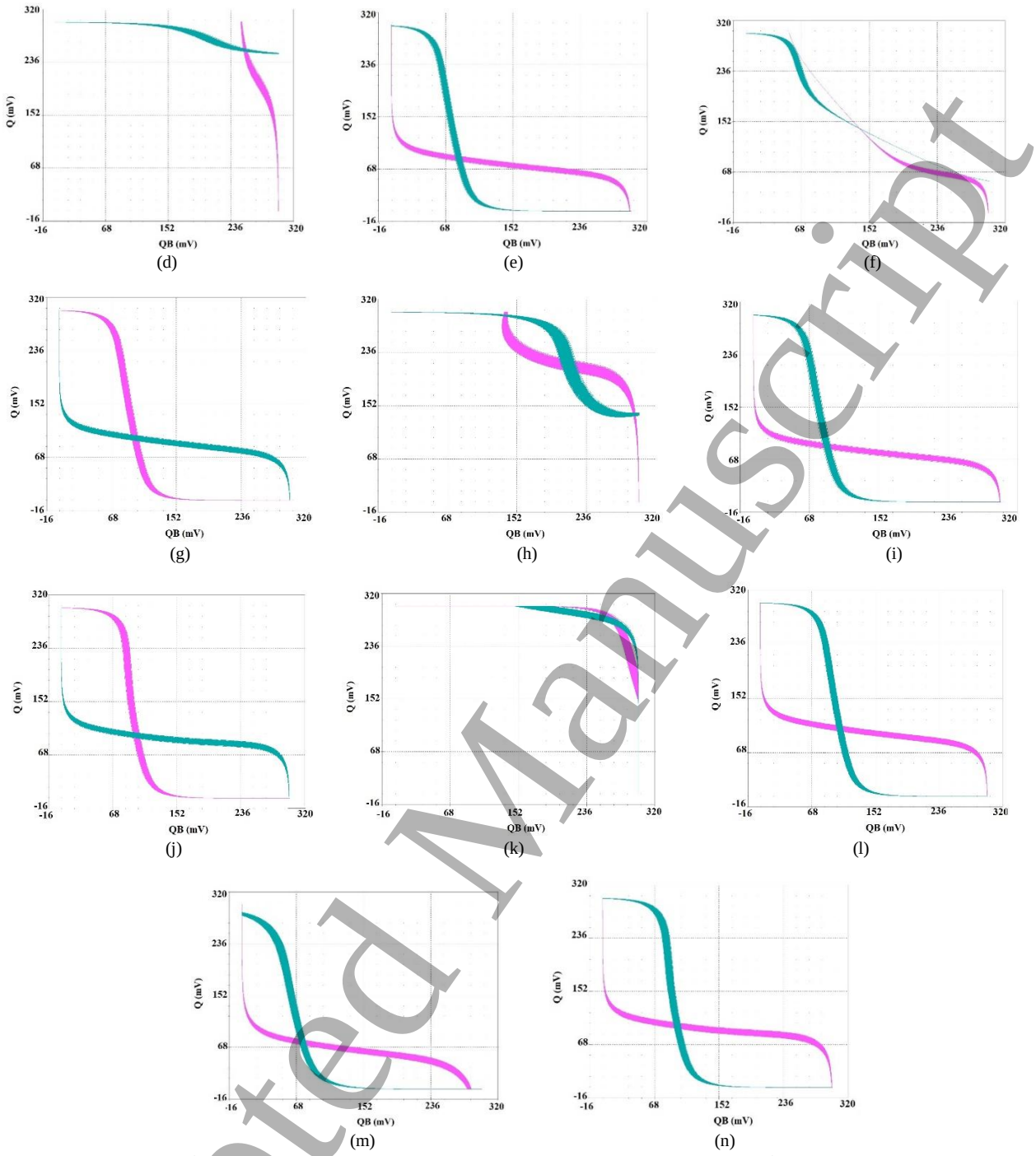
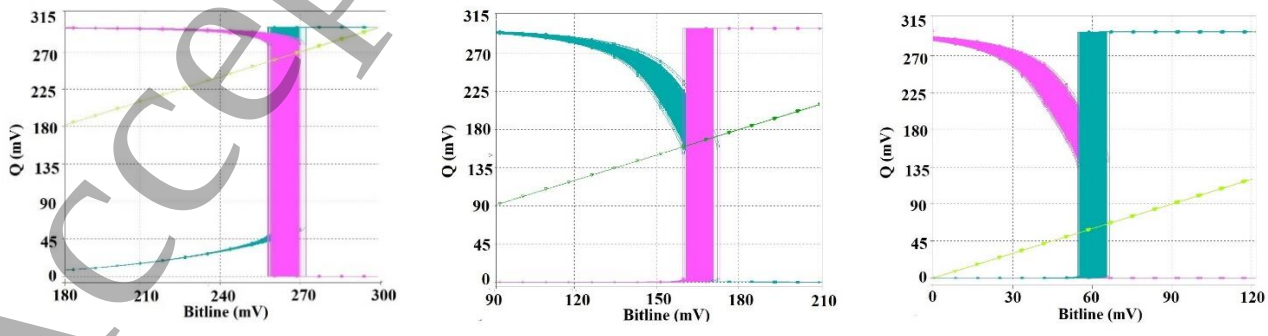


Fig. 6 Monte Carlo Simulations for (a) 5T HSNM (b) 5T RSNM, (c) 6T HSNM, (d) 6T RSNM, (e) 7T-1 HSNM, (f) 7T-1 RSNM, (g) 7T-2 SNM, (h) 7T-3 SNM, (i) 7T-4 SNM, (j) 7T-5 HSNM, (k) 7T-5 RSNM, (l) 8T SNM, (m) 9T SNM and (n) 10T SRAM bit cells



Accepted Manuscript

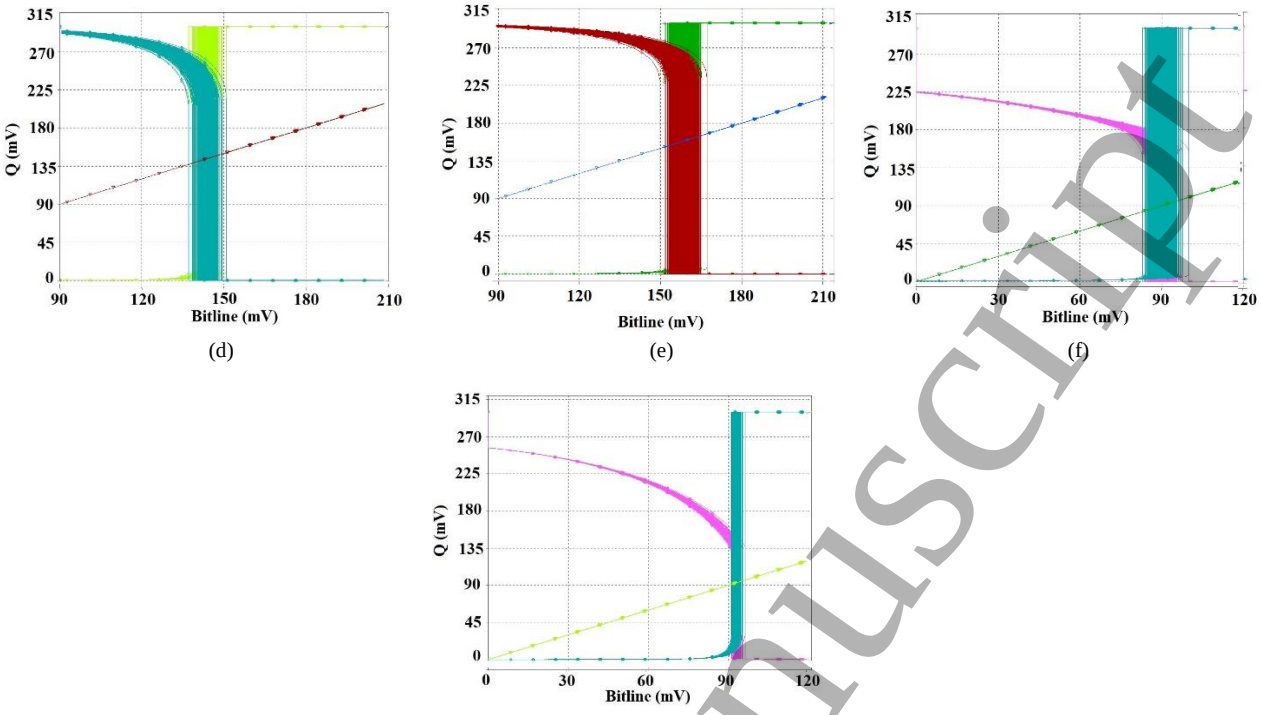


Fig. 7 Monte Carlo Simulations for WM for (a) 5T, (b) 6T, (c) 7T-1, (d) 7T-5, (e) 8T SNM, (f) 9T SNM and (g) 10T SRAM bit cells

TABLE II – The values obtained for minimum, maximum, mean, and standard deviation for all the bit cell at 32 nm technology node using Monte Carlo Analysis

Bit Cell	**HSNM (mV)**				**RSNM (mV)**				**WM (mV)**			
	Min	Max	Mean	S D	Min	Max	Mean	S D	Min	Max	Mean	S D
5T	66	88	77	11	BCC	BCC	-	-	30	60	45	15
6T	45	73	59	14	BCC	BCC	-	-	115	155	135	20
7T-1	64	84	74	10	BCC	BCC	-	-	230	250	240	10
7T-2	69	91	80	11	69	91	80	11	DP	DP	-	-
7T-3	21	43	32	11	21	43	32	11	DP	DP	-	-
7T-4	64	84	74	10	64	84	74	10	DP	DP	-	-
7T-5	73	89	81	8	4	18	11	7	120	160	140	20
8T	78	94	86	7.5	78	94	86	7.5	145	175	160	15
9T	45	73	59	14	45	73	59	14	180	230	205	25
10T	67	91	79	12	67	91	79	12	190	220	205	15
7TP	80	100	90	10	80	100	90	10	160	230	195	35

BCC – Butterfly Curve Collapsed, DP – Dual Pulse Write, SD – Standard Deviation

The cells that register a collapse of their butterfly for the read operation are 5T, 6T, 7T-1 and 7T-5. These cells have a poor RSNM perfromance as their read path discharges through the storage nodes of the memory core, making it highliy vulnerable to read error. While, the 7T-2, 7T-3, 7T-4, 8T, 9T and 10T SRAM bit cells have equivalent HSNM and RSNM values as these cells have an isolated read port that does not discharge through the storage node of the memory core. Since, HSNM and RSNM are equal, a single curve is presented for them. The Monte Carlo simulation for the WM for the SRAM bit cells are presented in Fig. 7.

The performance of the 7TP bit cell is compared against the other bit cells (with the process variation analysis). The results thus obatined for process variation analysis are tabulated in Table II. The Monte Carlo analysis yields the variation in static performance metric for the bit cells. Therefore, four values are tabulated corresponding to minimum, maximum, mean, and standard deviation (SD) for each performance parameter (i.e. HSNM, RSNM, and WM). The cells that register maximum variation in their performance, are the 6T and 9T SRAM bit cells. These two cells have a high deviation from their mean value. Additionally, the read static perfromance for the 6T cell at 300 mV is so inferior that the butterfly curve for RSNM is collapsed. A collapse in butterfly curve is also registered for 5T and 7T-1 cells. While, for 7T-5 SRAM bit cell the butterfly curve is on the verge of collapsing. It can be noted from Table II that the perfromance of the 7TP SRAM bit cell is well balanced as the Monte Carlo simulations for the bit cell are moderate for all three performance parameters.

Accepted Manuscript

Moving forwards the worst case analysis data is of higher concern, as its impact on the perfromance of the circuit is more deteriorative. Hence the data for the worst case (minimum values for each metric) is graphically compared in the Fig. 8. It is observed that the result for each bit cell deteriorate under the process variation analysis. The 5T, 6T, 7T-1 cell observe a significant reduction in HNSM, while the RSNM is reduced to zero. Thereby, ensuring read failure for the three cells. The WM for 5T reduces to 45 mV, which is exteremly low while that for 7T-1 increases to 240 mV (too high). The 6T has a comparable WM at 135 mV but due to its read failure, the cell is categorised as inferior to 7TP. The perfromane of 7T-3 is poor with respect to all static measures. The 7T-2 and 7T-4 cells have similar performance under process variation analysis. But owing to their dual write operation, the write perfroamcne cannot be compared. The Monte Carlo simulation HSNM for 7T-5 is comparable to 7TP, but its RSNM butterfly curve collapses, thereby making its perfromance inferior to the 7TP bit cell.

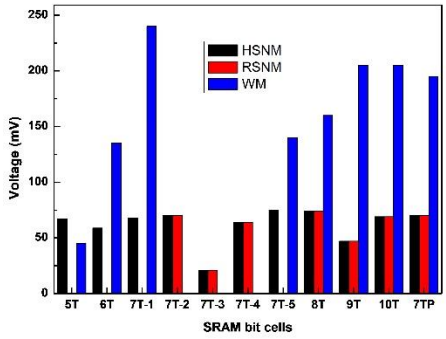


Fig. 8. Graphical comparison for the static margins obtained for all the SRAM bit cell under process variation analysis

Another cell that suffers and shows an inferior performance for process variation analysis is the 9T SRAM bit cell. Its HSNM and RSNM values reduce considerably, thereby making it less suitable for low voltage utilization. While the WM increases beyond 200 mV, making it extremly difficult to write into the cell. The 8T and 10T SRAM bit cells have a considerable performance with respect to HSNM and RSNM for the cell. But the WM for 8T is low, while that for 10T SRAM bit cell is very high at 116 and 230 mV respectively. The HSNM and RSNM for the proposed 7T SRAM bit cell is highest for the proposed 7T SRAM bit cell. Also its WM is fairly reasonable at 180 mV.

5.2 Temperature Variation Analysis

SRAM bit cells form a major part of SoC and these SoC's functions under wide temperature ranges. Due to continuous operation or uncontrolled environmental conditions, the device temperature may surge [58]. Therefore, an SRAM bit cell has to be analyzed for performance under varying environment temperature. So the performance of the 7TP SRAM bit cell is evalauted for temperature range of -10º to 80ºC.

This performance analysis ensures the reliability of the cell under varying environment temperature. The results for HSNM, RSNM and WM analysis for the 7TP SRAM bit cell are presented in Fig. 9. As can be observed from the graphs, the performance of the bit cell does show a bit of a deflection with temperature variation, but the performance of the cell is still within good range of performance. The bit cell is able to maintain HSNM and RSNM margin at 80mV, which is 11.1% deviated from the normal functioning. While the WM increases to 190 mV.

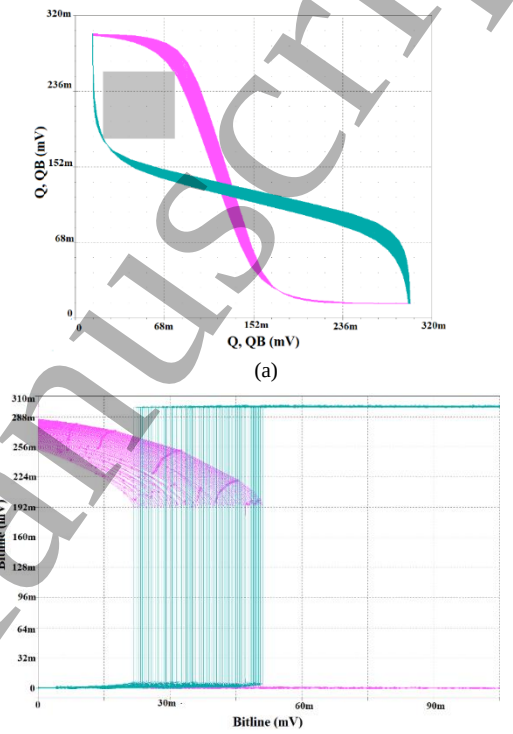


Fig. 9. Performance of the proposed 7T SRAM bit cell for temperature range from -10 C to 80 C (a) HSNM and RSNM and (b) WM

The read and hold margins are reduced by 0.1mV per degree Celsius. While the WM changes 0.2mV per degree Celsius. This data proves that the performance of the 7TP SRAM bit cell varies slightly with increase in temperature.

5.3 Voltage Variation Analysis

An SRAM cell may be subjected to varying values for V_{DD}. Also, it is imperative to analyse each cell for performance under sub-threshold and super threshold conditions. Therefore, static analysis for 7TP SRAM bit cell for voltage varying from 0.3 V up to 0.7 V is presented in this sub-section. The values obtained for HSNM/RSNM and WM are depicted in the Fig. 10. The graphical representation reflect that the 7TP SRAM bit cell maintains its superior performance for the entire voltage range.

Accepted Manuscript

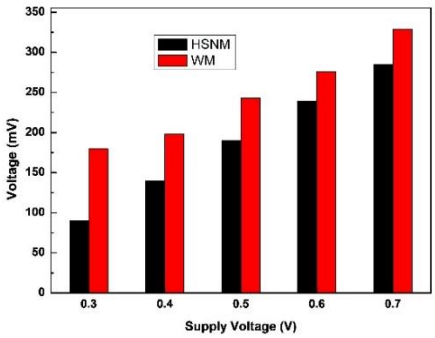


Fig. 10. Perofrmance of the proposed 7T SRAM bit cell with varying voltage

The operational V_{DD} increases, the SNM and the WM for the cell inceases. But the increase in the noise margin is more drastic for the sub-threshold region as compared to the super threshold region. As the increase of V_{DD} from 0.3 V to 0.4 V increases the HSNM by 55.5 %, while the increase is 35 % for 0.4 V to 0.5 V. The increase is 25.7 % and 19.2 % for 0.5 to 0.6 V and 0.6 to 0.7 V, respectively.

With the decreasing technology node, the vulnerability to error due to α-particles and cosmic neutrons is increasing. These particles generate extra charge through direct or indirect ionization in silicon, which collected by sensitive nodes, creates voltage transients at those nodes [59]. This transient caused in an SRAM bit cell causes an incorrectly stored datum and this error is referred to as soft error issue. The reduction in V_{DD}, adds to the vulnerability of SRAM bit cells towards soft error [60, 61]. Additionally, as explained by Jahinuzzaman *et al.* [62] in 2009 the process and temperature variations that a SRAM cell may encounter, add to the soft error susceptibility [63]. Therefore, the 7TP SRAM bit cell is analyzed for process and temperature variations. The analysis reveals that the cell is fairly tolerant to these variation and register a stable performance for the same. So in keeping with the reference [63] and conclusion stated by Jahinuzzaman *et al.* [62], we also conclude that our 7TP SRAM bit cell will demonstrate reasonable performance against soft error. Many researchers: Mansore *et al.* [64], Wen *et al.* [45], Shin *et al.* [46], Ahmad *et al.* [28] have also reported BI as an effective method to eliminate the soft error. BI technique is used to segregate the continuous bits of a single word by placing them separately [50]. This non-consecutive placement of bits on the same wordline to create an interleaving architecture is a preferred alternative to solve soft error issues. Researchers; Nayak *et al.* [65] have reported their 8T SRAM bit cell to be capable of soft error minimization based on BI. The same conclusion is also reported by researcher Pal and Islam [56] for their 9T SRAM bit cell reported in 2016. As BI architecture is supported by the 7TP SRAM bit cell, it should also demonstrate resilience to soft errors.

BI is a technique that places only one bit of a word at a particular location rather than all the bits of a word together. It is an efficient architectural technique to deal with soft error [50, 51]. Sachdeva and Tomar [50] in 2020 have reported an isolated read port based 10T SRAM bit cell. They have referred to BI as an efficient technique to shield their bit cell against soft errors. Similar information is also explained in the work reported by Mansore *et al.* [64] in 2019. The paper proposed an 11T SRAM bit cell, with a read SNM free isolated read port. However, BI technique is only applicable to the cells that exhibit a fully half-select free operation [28]. Using the same analogy for our 7TP SRAM bit cell, as it is also half select disturbance free, it can be stated that utilizing BI will make the 7TP SRAM bit cell resilient to soft errors. BI is also referred to as column selection architecture. SRAM bit cells employ this technique to achieve area efficiency and provide soft error protection as reported by Wen *et al.* [45] in 2016. This work also suggests that cells with half select disturbance free architecture are appropriate for BI configuration. Thereby supporting our claim that as the proposed 7T SRAM bit cell has a fully half-select free operation, utilizing bit-interleaving for designing the SRAM array will make the cell less vulnerable to errors. The cell is half select disturbance free due to its row and column based independent signals for read and write operations.

6. Failure Probability

The stability for the bit cell is estimated based on the HSNM, RSNM and the WM values for the bit cell. Whereas, the failure probability for the three operations: read, hold and write, are estimated using eq. (1-3) [66]

$$\boldsymbol{P_{R-Fail} = Prob.(RSNM < kT)} \quad (1)$$

$$\boldsymbol{P_{H-Fail} = Prob.(RSNM < kT)} \text{ and} \quad (2)$$

$$\boldsymbol{P_{W-Fail} = Prob.(WM < 0\ mV)} \quad (3)$$

If, the RSNM and HSNM are lower than thermal voltage (kT = 26 mV at 300 K), the thermal noise will cause a bit error in the SRAM cell. The bit error rate (BER) is defined as the probability of a random event causing an error in the bit cell during its operation [67]. The failure probabilities for each cell are calculated and the obtained values are presented in Fig. 11.

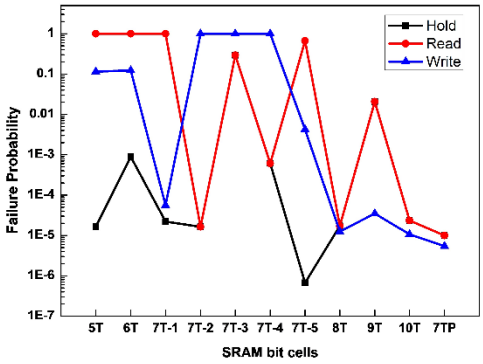


Fig. 11. Failure Probability Comparison for all SRAM bit cell for read, hold and write operation.

In keeping with the process variation analysis, the read failure probability for 5T, 6T, 7T-1 is unity, as they register a zero RSNM under process variation analysis. Similarly, the write failure probability for 7T-3 and 7T-4 are also unity, as they have zero WM. The 7T-2 cell is shown to have unity

Accepted Manuscript

failure probability because its WM cannot be realized as it has a dual cycle write operation. The failure probablity for 7T-5 a contrasting as its HSNM failure probablity is very low due to its high HSNM. While for RSNM the failure probability is high, owing to its differential nature and its WM failure probablity is also moderatly high. WM. The 8T, 9T and 10T SRAM bit cells have lower failuer probablity in comparison to the other SRAM bit cells. The lower deviation from the static margin for the bit cell leads to the lower failure probability. The 7TP SRAM bit cell registers the least failure probability, thereby making it more BER tolerate than other bit cells.

As discussed in the previous subsection, another issue that arises in SRAM bit cell as the V_{DD} and technology node are shrinking is the soft errors [26]. These vulnerability are applicable to the 7TP SRAM bit cell as well. But as the tolerance variability of the 7TP SRAM bit cell is high and also its failure probability being low; the cell can be infreed to be more reliable than the other cells. Also, since the 7TP SRAM bit cell is half select disturbance free, employing BI makes the 7TP SRAM memory tolerant to soft errors. Also, employing interleaving of multiple word lines onto same physical row [68] would help to avoid soft error concerns.

7. I_{ON}/I_{OFF} Analysis

The relationship between the read current and leakage current is called the I_{ON}/I_{OFF} ratio. Greater I_{ON}/I_{OFF} ratio allows the designer to create larger SRAM array. A column in the SRAM array shares a sense amplifier for the read circuit. So, a greater current ratio ensures that a designer can integrate greater number of bit cells per column. Thus, reducing the bulk of read peripheral circuit in the memory architecture. Therefore, a bit cell with higher ratio is preferred over other bit cells. The comparison for the I_{ON}/I_{OFF} ratio for all the SRAM bit cells is plotted in Fig. 12.

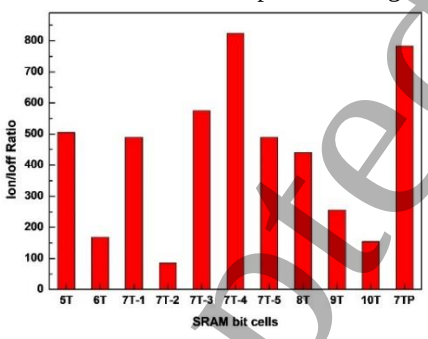
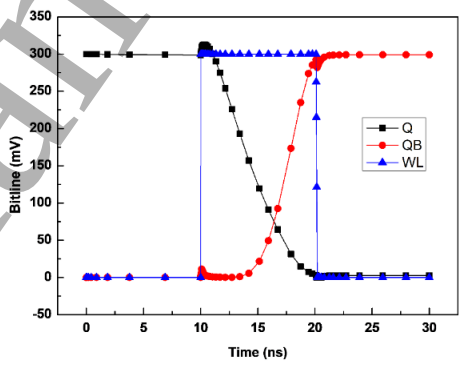


Fig. 12. Comparison of the I_{ON}/I_{OFF} ratio for all the SRAM bit cells

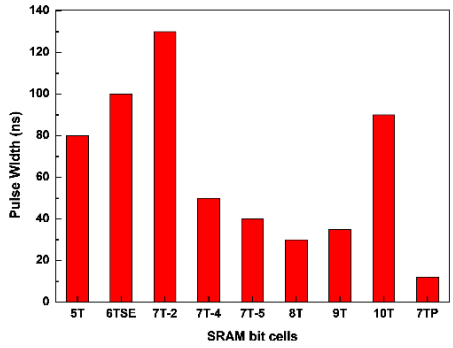
On the basis of the graphical comparison for I_{ON}/I_{OFF} presented in Fig. 12, it can be concluded that the 7TP SRAM bit cell has improved performance with respect to I_{ON}/I_{OFF} ratio. The bit cells 6T, 7T-2 and 10T have a poor performance, as their I_{ON}/I_{OFF} ratios are 168, 85 and 154 respectively. While the 5T, 7T-1, 7T-3, 7T-5, 8T, and 9T SRAM bit cells display moderate performance with the ratio being 505, 489, 575, 489, 440, and 256 respectively. The two cells that have the highest ratio values are 7T-4 and 7TP at 824 and 783, respectively. Therefore, for hardware implementation they would be the preferred choice. The bit cells 5T, 7T-1 and 8T show a reasonable performance but the 7TP SRAM bit cell out ranks them under this evaluation. The 9T and 10T SRAM bit cells suffer highly in the current ratio metric because of their stacked inverter configuration.

8. Dynamic Write Analysis

The static analysis of WM assumes infinite pulse width for the write operation. While, real time write operations for an SRAM bit cell is dependent on a finite pulse width. Therefore, determining the minimum pulse width for a successful read operation is imperative for an SRAM bit cell design. The minimum pulse width needed to perform a successful write operation is referred to as write pulse width (T_{crit}). In this work the T_{crit} is estimated using the method explained in reference [42]. If, the pulse width on WL is lower than T_{crit}, it leads to a write failure for that cycle. Therefore, it is imperative to ensure that the pulse width of WL is greater than T_{crit}. A successful write operation at 300 mV for the 7TP SRAM bit cell is shown in Fig. 13 (a). The T_{crit} for the proposed cell is found to be 10 ns at 27 ºC.



(a)



(b)

Fig. 13. (a) Dynamic write operation for finite pulse width of 10 ns for the proposed 7TP SRAM bit cell and (b) Graphical comparison for the T_{crit} needed for successful write operation for the different SRAM bit cells.

The pulse width needed for the other SRAM bit cells to perform a successful write operation are compared in Fig. 13 (b). All the bit cells require pulse widths multi-folds larger than the 7TP SRAM bit cell. The 5T, 6T, 7T-2, 7T-4, 7T-5 8T, 9T and 10T SRAM bit cell require the pulse width to be 80, 100, 130, 50, 40, 30, 35 and 90 ns respectively. The 7T-3 SRAM fails to perform dynamic write operation at 300 mV V_{DD}, whereas the 7T-1 cell is not included in the dynamic write analysis as it has a two-step write process. Therefore,

Accepted Manuscript

the comparison is avoided, to circumvent projecting it as having false poor performance in single pulse analysis.

9. Power Consumption Analysis

SRAM occupies a major section of SoC and thus, the total power consumption by the system will have a major contribution from the memory circuit. Therefore, analyzing the power consumption of the SRAM bit cell during each of its operations is important. The current requirements for the bit cells are different for the three operations – read, write and hold. Also, the power consumption of the circuit depend on the value stored in the bit cell. So, while evaluating the power consumption for each operation, the power is computed for X = '0' and X = '1'. Thereby, providing a very comprehensive understanding of the power consumption of the bit cell.

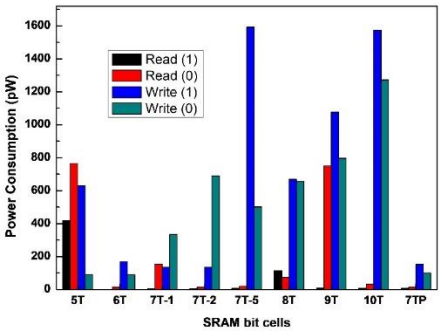


Fig. 14. Comparison of read and write power consumption for various SRAM bit cells

Power consumption in an SRAM bit cell is maximum during the read and write operations, as the memory core of the bit cell is being accessed. Amongst the two, higher power consumption takes place during the write operation. As the data stored in the core of the bit cell has to be flipped to perform a successful write operation. While, the read operation requires the voltage on the bit line to drop to a point that the data in the memory core of the bit cell can be sensed by the sense amplifier. Therefore, the read power consumption for the bit cells is lower as compared to write power consumption. The graphical comparison of the read and write power of the bit cell for X = '0' and X = '1' are given in Fig. 14. It can be observed from the power consumption graphs, that the 7TP SRAM bit cell has the least power consumption for read as well as write operation. This improved performance of the 7TP SRAM bit cell is credited to the optimal sizing of the transistors and the presence of a LV_{TH} NMOS.

The 7TP SRAM bit cell relies on reduction of α_T to reduce the dynamic power consumption [69]. The 7TP bit cell utilizes the same and is dependent on a single bit line architecture for its operation. This single bitline structure also helps in reducing the C_L for the cell, thereby ensuring a three-fold reduction in the dynamic power consumption of the 7TP SRAM bit cell. Additionally, as the 7TP cell has a considerably low τ , it can be justified that the 7TP SRAM is bound to have low short circuit power.

Another major factor that contributes to the total power consumption of the bit cell is static power consumption. Leakage power is the major contributor to static power consumption. SRAM bit cell primarily operates in the hold mode. Therefore, power consumption during the hold mode – referred to as standby power - forms a major component of the total leakage power consumption [47]. Hence, it is imperative analyze the standby power for SRAM bit cell. Additionally, in submicron technologies, standby power dissipation is the major component of overall power consumption and can be attributed to the increased leakage current of nano-scaled device.

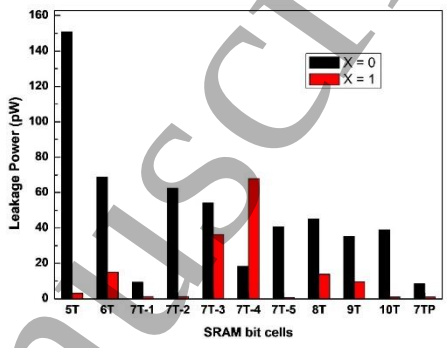


Fig. 15. Comparison for standby (hold) power for all SRAM bit cells.

To evaluate the leakage current, the cell is set in the hold operation mode, i.e. the data values are stored in the cell. Then the access transistors are turned OFF. Once the access transistors are OFF the current at the drain terminal for the transistor is measured. During the process, the bit line is set high and then corresponding to data stored at X = '0' and X = '1', the leakage current is calculated for the bit cells. The leakage power is calculated V_{DD} times the leakage current for all the bit cells. The comparison for the leakage power for all SRAM bit cells discussed so far is presented in Fig. 15.
The above bar graph confirm that the proposed SRAM bit cell has the least standby power for both X = '0' and '1' at 8.4 and 1.2 pW, respectively. The standby power for X = '0' is highest for 5T bit cell while for X = '1', the highest standby power is recorded for 6T SRAM bit cell. This hold mode power at X = '0' for 5T, 6T, 7T-1, 7T-2, 7T-3, 7T-4, 7T-5, 8T, 9T and 10T bit cell is greater by 16.9, 7.17, 0.11, 6.42, 5.46, 1.18, 3.82, 4.35, 3.17 and 3.64 than the leakage power for 7TP cell. While for X = '1' the standby power is greater by 1.85, 13.28, 0.14, 0.17, 33.57, 63.57, 0.15, 12.14, 8.14 and 0.14 times in comparison to the 7TP bit cell.

10. Layout and Area Analysis

Area analysis is imperative as SRAM bit cells are integrated to form large cache memory. This memory occupies a significant area on the SoC. The layouts for all the SRAM bit cells are presented in Fig. 16. The dimensions for the layout are marked on the layout itself. For the ease of comparison, the length (µm), breadth (µm) and area (µm^2) are tabulated in Table III.
It can be observed from the layout that the 7TP SRAM bit cell has the second lowest area footprint. The only other

Accepted Manuscript

SRAM bit cell with area less than 7TP SRAM cell is the 7T-3 bit cell. But this cell as demonstrated in the previous sections has a considerable poor performance. This validates that even if the transistor count increases, the area may turn out to be less if the transistors are sized accurately. Similar results were reported by Pal *et al.* [56] in 2020, wherein, their proposed 9T SRAM bit cell exhibited area 7% less than the pre-existing 8T SRAM bit cell. Gupta *et al.* [70] in 2018 have also reported a 10T SRAM bit cell with area 7% less in comparison to a pre-existing 8T SRAM bit cell.

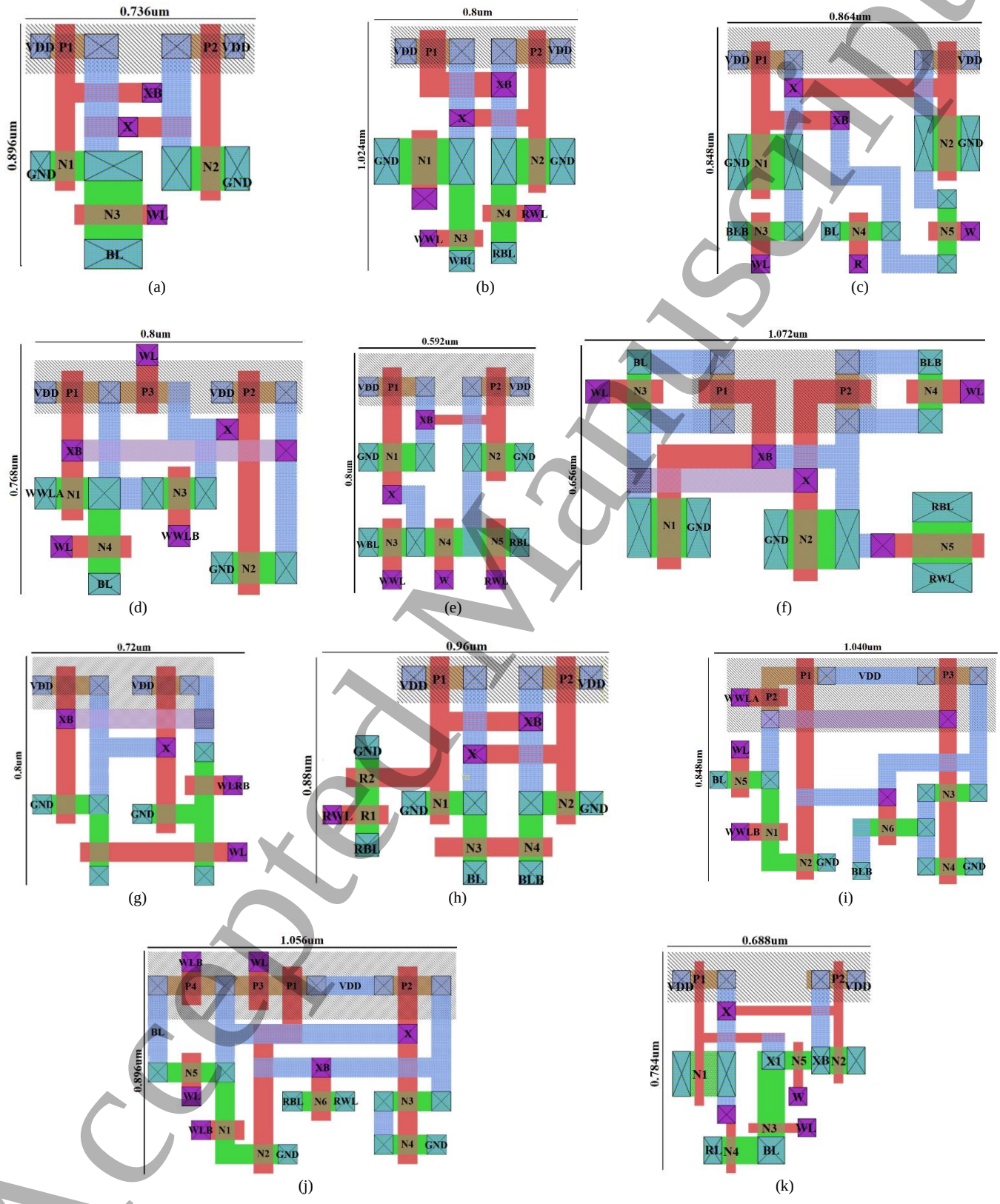


Fig. 16. Layouts for (a) 5T, (b) 6T, (c) 7T-1, (d) 7T-2, (e) 7T-3, (f) 7T-4, (g) 7T-5, (h) 8T, (i) 9T, (j) 10T and (k) 7TP SRAM bit cells

Accepted Manuscript

TABLE III Layout dimensions for the SRAM bit cells

SRAM bit cells	L (μm)	W (μm)	Area (μm^2)
5T	0.736	0.896	0.659
6T	0.8	1.024	0.819
7T-1	0.864	0.848	0.733
7T-2	0.8	0.768	0.614
7T-3	0.592	0.8	0.474
7T-4	1.072	0.656	0.703
7T-5	0.72	0.8	0.576
8T	0.96	0.88	0.845
9T	1.040	0.848	0.882
10T	1.056	0.896	0.946
7TP	**0.688**	**0.784**	**0.539**

The area of the proposed SRAM bit cell is 0.539 μm^2. The area for 5T, 6T, 7T-1, 7T-2, 7T-4, 8T, 9T and 10T SRAM bit cells are 0.659, 0.819, 0.614, 0.733, 0.703, 0.845, 0.882 and 0.946 μm^2 respectively. With reference to 7TP bit cell, 5T, 6T, 7T-1, 7T-2, 7T-4, 7T-5, 8T, 9T and 10T cells occupy more area by 22.17, 51.8, 35.8, 13.9, 30.4, 6.78, 56.6, 63.6 and 75.5 %.

11. Conclusion

In this paper a single port, seven transistor SRAM bit cell is proposed. The architecture of the bit cell has a single bit line configuration. The proposed 7T SRAM bit cell is reported for 32 nm technology node. The bit cell is compared with other SRAM bit cells that have single ended operation. The bit cell is operational at 300 mV and maintains HSNM, RSNM and WM at 90, 90 and 180 mV respectively. The cell has the most balanced static margin in comparison to the other single ended SRAM bit cells that are referred in the paper. The Monte Carlo analysis of the bit cell highlights that even under *6σ* global variation, the cell maintains 75 mV as the read as well as hold SNM, while the WM is 215mV. This validates the stability of the bit cell under process variation. The temperature variation analysis for temperature ranging from -10ºC to 80ºC reveals that HSNM and RSNM are reduced by 0.1mV per degree Celsius. While the WM changes 0.2mV per degree Celsius. The pulse width needed for the cell to perform a successful write opration in 10 ns. While the power analysis portrays that the proposed SRAM bit cell maintains a low power consumption of all operation. The highlight being the least standby power of 8.4 and 1.05 pW for X = '0' and '1' respectively. These aforementioned parameters are maintained with a very high I_{ON}/I_{OFF} ratio of 783. This implies that the number of bit cell per column of the 7TP SRAM bit cell is higher in comparison to other bit cells. The layout for the bit cell occupies 0.539 μm^2 area. This area footprint for the bit cell is least amongst all the other cells in the study. The 5T, 6T, 7T-1, 7T-2, 7T-4, 8T, 9T and 10T SRAM bit cells have area greater than the proposed cell by 22.17, 51.8, 35.8, 13.9, 30.4, 6.78, 56.6, 63.6 and 75.5 %. Thus, it is valid to conclude that the 7TP SRAM bit cell showcases an all-inclusive improvement in performance when compared with other single ended counterparts discussed in this paper.

References

[1] Pourbakhsh SA, Chen X, Chen D, Wang X, Gong N and Wang J 2016 Sizing-priority based low-power embedded memory for mobile video applications *2016 17th International Symposium on Quality Electronic Design (ISQED)*.

[2] Semiconductor Industry Association (SIA), International technology roadmap for semiconductors, 2013 edition.

[3] Yang B and Kim L 2005 A low-power SRAM using hierarchical bit line and local sense amplifiers *IEEE Journal Solid State Circuit* **40** 1366.

[4] Hobson RF 2007 A new single-ended SRAM cell with write-assist *IEEE Transaction on Very Large Scale Integration (VLSI) Systems* **15** 173.

[5] Autran JL, Munteanu D, Moindjie S, Saoud TS, Gasiot G and Roche P 2016 Real-time soft error rate measurements on bulk 40 nm SRAM memories: a five-year dual-site experiment *Semiconductor Science and Technology*, **31** 114003.

[6] Sil A, Ghosh S and Bayoumi MA 2007 A novel 90 nm 8T SRAM cell with enhanced stability *IEEE International. Conference Integration Circuit Des.Technology*

[7] Noguchi H, Okumura S, Iguchi Y, Fujiwara H, Morita Y, Nii K, Kawaguchi H and Yoshimoto M 2007 Which is the best dual-port SRAM in 45-nm process technology? 8 T, 10 T single end, and 10 T differential *Proc. ICICDT* 55.

[8] Calhoun BH, Ryan JF, Khanna S, Putic M and Lach J 2010 Flexible circuits and architectures for ultra low power *Journal of. Flexible circuits and architectures for ultralow power. Proc. IEEE*, **98** 267.

[9] Sharma V, Vishvakarma S, Chouhan SS and Halonen K 2018 A write improved low power 12T SRAM cell for wearable wireless sensor nodes *International Journal of Circuit Theory Application* **46** 2314.

[10] Cho K, Park J, Oh TW and Jung OK 2020 One sided Schmitt-Trigger Based 9T SRAM cell for near threshold operation *IEEE Transactions on Circuits and Systems I: Regular Papers* **67** 1551.

[11] Mishra N, Mittal P and Kumar B 2019 Analytical Modeling for Static and Dynamic Response of Organic Pseudo All-p Inverter Circuits *Journal of Computational Electronics* **18** 1490.

[12] J. P. Kulkarni, K. Kim and K. Roy, "A 160 mV Robust Schmitt Trigger Based Subthreshold SRAM," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 10, pp. 2303–2313, 2007.

[13] Mittal P, Negi YS and Singh RK 2016 A Depth Analysis for Different Structures of Organic Thin Film Transistors: Modeling of Performance Limiting Issues *Microelectronics Engineering* **150** 7.

[14] Kumar M and Ubhi JS 2019 Design and analysis of CNTFET based 10T SRAM for high performance at nanoscale *International Journal of Circuit Theory Applications* **47,** 1775.

[15] Faraji R, Naji HR, Nezhad MR and Arabnejhad M 2014 A new SRAM design using body bias technique for low power and high speed application *International Journal of Circuit Theory Applications* **42** 1189.

Accepted Manuscript

[16] Toh SO, Guo Z, Liu TJK and Nikolic B 2011 Characterization of dynamic SRAM stability in 45nm CMOS *IEEE J. Solid-State Circuits* **46** 2702.

[17] Calhoun BH and Chandrakasan AP 2006 Static noise margin variation for subthreshold SRAM in 65-nm CMOS *IEEE Journal Solid-State Circuits* **41** 1673.

[18] Karl E, Wang Y, Ng YG, Guo Z, Hamzaoglu F, Meterelliyoz M, Keane, Bhattacharya U, Zhang, Mistry K and Bohr M, 2013 A 4.6 GHz 162 Mb SRAM design in 22 nm Tri-gate CMOS technology with integrated read and write assist circuitry *IEEE J. Solid-State Circuits* **48** 150.

[19] T. Song, Rim W, Jung J, Yang G, Park J, Park S, Kim Y, Baek KH, Beak S, Oh SK, Jung J, Kim S, Kim G, Kim J, Lee Y, Sim SP, Yoon JS and Choi KM 2014 A 14 nm FinFET 128 Mb SRAM with V*MIN* enhancement techniques for low-power applications *IEEE J. Solid- State Circuits* **50** 158.

[20] Yabuuchi M, Nii K, Tsukamoto Y, Ohbayashi S, Nakase Y and Shinohara H 2009 A 45 nm 0.6 V cross-point 8T SRAM with negative biased read/write assist *Proc. Symp. VLSI Circuits*, 2009, 158.

[21] Chang J, Chen YH, Cheng H, Chan WM, Liao HJ, Li Q, Chang S, Natarajan S, Lee R, Wang PW, Lin SS, Wu CC, Chng KL, Cao M and Chang GH 2013 A 20 nm 112 Mb SRAM in High-*k* metal-gate with assist circuitry for low-leakage and low-V*MIN* applications *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers (ISSCC)* 316.

[22] Raychowdhury A, Geuskens B, Kulkarni J, Tschanz J, Bowman K, Karnik T, Lu SL, De V and Khellah MM 2010 PVT-and-aging adaptive wordline boosting for 8T SRAM power reduction *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers (ISSCC)* 352.

[23] Chang L, Fried DM, Hergenrother J, Sleight JW, Dennard RH, Montoye RK and Guarini KW 2005 Stable SRAM cell design for the 32 nm node and beyond. In Digest of Technical Papers *2005 Symposium on VLSI Technology* 128.

[24] Ahmad S, Ahmad SA, Muqeem M, Alam N and Hasan M 2019 TFET-Based Robust 7T SRAM cell for Low Power Appplication *IEEE Transactions in Electron Devices* **66**(9), 3834-3840.

[25] Wen L, Li Z, and Li Y 2013 Single-ended robust 8T SRAM cell for low voltage operations *Microelectron. Journal* **44**(8) 718.

[26] Sil A, Ghosh S, Gogineni N and Bayoumi M 2008 A Novel High Write Speed, Low Power, Read-SNM-Free 6T SRAM Cell *2008 51st Midwest Symposium on Circuits and Systems.*

[27] He Y, Zhang J, Wu X, Zhen S and Zhang Bo 2019 A Half-Select Disturb-Free 11T SRAM Cell With Built-In Write/Read-Assist Scheme for Ultralow-Voltage Operations *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* **27**(10), 2344.

[28] Ahmad S, Iqbal B, Alam N and Hasan M 2018 Low Leakage Fully Half-Select-Free Robust SRAM Cells with BTI Reliability Analysis *IEEE Transactions on Devices and Material Reliability* **18**(3) 337.

[29] Chang IJ, Kim JJ, Park S and Roy K 2009 A 32 kb 10T sub-threshold SRAM array with bit-interleaving and differential read scheme in 90 nm CMOM *IEEE Journal of Solid State Circuits* **44**(2) 650.

[30] Carlson I, Andersson S, Natarajan S and Alvandpour A 2004 A high density, low leakage, 5T SRAM for embedded caches *Proceedings 30th European Solid State Circuits Conference*, 215

[31] Surana N and Mekie J 2019 Energy Efficient Single-Ended 6-T SRAM for Multimedia Applications *IEEE Transactions on Circuits and Systems II: Express Briefs* **66** 1023.

[32] Aly RE and Bayoumi MA 2007 Low-Power Cache Design Using 7T SRAM Cell *IEEE Transactions on Circuits and Systems II: Express Briefs* **54** 318.

[33] Yang Y, Jeong H, Song SC, Wang J, Yeap G and Jung SO 2016 Single Bit-Line 7T SRAM Cell for Near-Threshold Voltage Operation With Enhanced Performance and Energy in 14 nm FinFET Technology *IEEE Transactions on Circuits and Systems I: Regular Papers* **63** 1023.

[34] Sanapala K, R S and Yeo SS 2018 Schmitt trigger-based single-ended 7T SRAM cell for Internet of Things (IoT) aaplications *Journal of Supercomputing* **74**, 4613-4622.

[35] Liu JS, Clavel MB and Hudait MK 2017 An energy-efficient tensile strained Ge/InGaAs TFET 7T SRAM cell architecture for ultra-low voltage applications *IEEE Trans. Electron Devices* **64**(5) 2193–2200.

[36] Oh JS, Park J, Cho K, Oh TW and Jung SO 2021 Differential Read/Write 7T SRAM with Bit-Interleaved Structure for Near-Threshold Operation *IEEE Access* **9**, 64104.

[37] Chang L, Montoye RK, Nakamura Y, Batson KA, Eickemeyer RJ Dennard RH, Haensch and Jamsek D 2008 An 8T-SRAM for variability tolerance and low-voltage operation in high-performance caches *IEEE Journal of Solid-State Circuits* **43**(4) 956.

[38] Lin S, Kim YB and Lombardi F 2008 A highly-stable nanometer memory for low power design *Proc. IEEE International Workshop Design Test of Nano Devices, Circuits Syst.* 17.

[39] Cho K, Park J, Oh TW and Jung SO 2020 One-Sided Schmitt-Trigger-Based 9T SRAM Cell for Near-Threshold Operation *IEEE Transactions on Circuits Systems I: Regular Papers* **67**(5) 1551.

[40] Eslami N, Ebrahimi B, Shakouri E and Najafi D 2020, A single-ended low leakage and low voltage 10T SRAM cell with high yield *Analog Integrated Circuits and Signal Processing* **105** 263.

[41] Chang MF, Wu JJ, Chen KT, Chen YC, Chen YH, Lee R, Liao HJ and Yamauchi H 2010 A differential data-aware power-supplied (D2AP) 8T SRAM cell with expanded write/read stabilities for lower VDDmin applications *IEEE Journal of Solid-State Circuits* **45** 1234.

[42] SinghJ, Mohanty SP and Pradhan DK Robust SRAM designs and analysis. New York: Springer, 2013.

[43] Kushwah CB and Vishvakarma SK 2016 A single-ended with dynamic feedback control 8T subthreshold SRAM cell *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* **24**(1) 373.

[44] Pasandi G and Fakhraie SM 2014 An 8T Low-Voltage and Low-Leakage Half-Selection Disturb-Free SRAM using bulk-CMOS and FinFETs *IEEE Transactions on Electron Devices* **61**(7) 2357.

[45] Wen L, Cheng X, Zhou K Tian and Zeng X Bit-Interleaving-Enabled 8T SRAM with Shared Data-Aware-Write and Reference-Based Sense Amplifier *IEEE Transactions on Circuits and Systems II: Express Briefs* **63**(7) 643.

[46] Shin K, Choi W and Park J 2017 Half-Select Free and Bit-Line Sharing 9T SRAM for Reliable Supply Voltage Scaling *IEEE Transactions on Circuits and Systems I: Regular Papers* **64**(8) 2036.

Accepted Manuscript

[47] Ahmad S, Gupta MK, Alam N and Hasan M 2016 Single-ended Schmitt Trigger based robust low power SRAM cell *IEEE, Transaction on Very Large Scale Integration (VLSI) Systems* **20** 2634.

[48] Ding J and Asenov A 2020 The analysis of static random access memory stability under the influence of statistical variability and bias temperature instability-induced ageing *Semiconductor Science and Technology* **36** 025008.

[49] Rawat B, Mittal P 2021 Single bit line accessed high high performance ultra low voltage operating 7T SRAM bit cell with improved read stability *International Journal of Circuit Theory and Application*, **49**(5) 1435.

[50] Sachdeva A and Tomar VK 2021 Design of Low Power Half Select Free 10T Static Random Access Memory cell *Journal of Circutis, Systems and Computers* **30**(4)**.**

[51] Chiu YW, Hu YH, Tu MH, Zhao JK, Chu YH, Jou SJ and Chuang CT 2014 40nm bit-interleaving 12T subthreshold SRAM with data-aware write-assist *IEEE Transactions on Circuits Systems I: Regular Papers* **61** 2578.

[52] Jin Q, Li X and Bernstein JB 2007 SRAM Stability Analysis Considering Gate Oxide SBD, NBTI and HCI *2007 IEEE International Integrated Reliability Workshop Final Report*

[53] Shah AP, Yadav N, Beohar A and Vishvakarma SK 2007 On-chip Adaptive Body Bias for Reducing the Impact of NBTI on 6T SRAM cells *IEEE Transactions on Semiconsuctor Manufacturing* **31**(2) 242.

[54] Agarwal A, Paul BC, Mukhopadhyay S and Roy K 2005 Process variation in embedded memories: failure analysis and variation aware architecture *IEEE Journal of Solid-State Circuits* **40**(9)1804.

[55] Wang Y, Cotofana S and Fang L 2011 A unified aging model of NBTI nd HCI degradation towards lifetime reliability management for nanoscale MOSFET circuits *2011 IEEE/ACM International Symposium on NanoscaleArchitectures*, 175.

[56] Pal S and Islam A 2016 9T SRAM Cell for Reliable Ultralow Power Applications and Solving Multi-Bit Soft Error Issue *IEEE Transactions on Device and Material Reliability* **16**(2) 172.

[57] Doorn T, Maten ET, Croon J, Bucchianico AD and Wittich O 2008 Importance sampling Monte Carlo simulations for accurate estimation of SRAM yield *ESSCIRC 2008 - 34th European Solid-State Circuits Conference*.

[58] Rawat B and Mittal P 2021 Analysis of varied architectural configuration for 7T SRAM bit cell. *4th International Conference on Recent Trend in Communication & Electronics (ICCE-2020)*

[59] Dodd PE and Massengill LW 2003 Basic mechanisms and modeling of single-event upset in digital microelectronics *IEEE Transactions on Device Material and Reliability* **50**(3) 583.

[60] Baumann R 2005 Soft errors in advanced computer systems *IEEE Design and Test of Computers* **22** 258.

[61] Granlund T, Granbom B and Olsson N 2003 Soft error rate increase for new generations of SRAMs *IEEE Transactions on Nuclear Sc*iences **50**(6) 2065.

[62] Jahinuzzaman SH, Rennie DJ and Sachdev M 2009 A Soft Error Tolerant 10T SRAM Bit-Cell With Differential Capability *IEEE Transactions on Nuclear Sc*iences **56**(6) 3768.

[63] Jahinuzzaman SM, Sharifkhani M and Sachdev M 2009 An analytical model for soft error critical charge of nanometric srams *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* **17**(9) 1187.

[64] Mansore SR, Gamad RS and Mishra DK 2020 A 32 nm Read Disturb-free 11T SRAM Cell with Improved Write Ability *Journal of Circuits, Systems and Computers* **29**(5) 2050067.

[65] Nayak D, Acharya DP, Rout PK and Nanda U 2018 A hihg stable 8T-SRAM with bit interleaving capability for minimization of soft error rate *Microelectronics Journal* **73** 43.

[66] Kulkarni JP and Roy K 2012 Ultralow-Voltage Process-Variation-Tolerant Schmitt-Trigger-Based SRAM Design *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*,**20**(2) 319–332.

[67] Suresh VB and Kundu S 2013 On analysis and mitigating SRAM BER due to rando thermal noise, *2013 IEEE Computer Society Annual Symposum on VLSI (ISVLSI)*, 159-164.

[68] Suzuki T, Yamagami Y, Hatanaka I, Shibayama A, Akamatsu H and Ymauchi H 2006 A sub-0.5-v operating embedded SRAM featuring a multi-bit-error-immune hidden-ecc scheme. *IEEE Journal of Solid-State Circuits* **41**(1) 152–160.

[69] Kumar N and Mittal P 2020 Performance Analysis of FinFET based 2:1 Multiplexers for low power applications *6th Students' Conference on Engineering and Systems (SCES-2020)*

[70] Gupta S, Gupta K and Pandey N 2018 Pentavariate Vmin Analysis of a Subthreshold 10T SRAM Bit Cell With Variation Tolerant Write and Divided Bit-Line Read *IEEE Transaction on Circuits and Systems I: Regular Papers* **65**(10) 3326.

Accepted Manuscript

Materials Today: Proceedings xxx (xxxx) xxx

Contents lists available at ScienceDirect

Materials Today: Proceedings

journal homepage: www.elsevier.com/locate/matpr

A *first principles* based study of the effect of uniform and tetragonal strains on half-metallicity in FeCrAs Heusler alloy

Aniket Singh, Mansher Singh, Mukhtiyar Singh *

Department of Applied Physics, Delhi Technological University, Delhi-42, India

ARTICLE INFO

Article history:
Received 19 April 2021
Received in revised form 19 May 2021
Accepted 26 May 2021
Available online xxxx

Keywords:
Spintronics
Heusler alloys
Half-metallic ferromagnetism
DFT

ABSTRACT

Half-Heusler alloys are prominent material for spintronics device as these materials usually exhibit high spin polarization. But this spin polarization is severely affected by various disorders and distortions which ultimately hinders their practical applications. This paper examines the effect of the uniaxial strain as well as the tetragonal strain on the electronic and magnetic properties of half-Heusler FeCrAs alloy. This work is carried out by using the first-principles density functional theory calculations using the Wein2k software. The results of our study show that FeCrAs alloy exhibiting AFM and FM characteristics in u- strain while showing only FM characteristics in t- strain. The Half metallicity of the FeCrAs is sensitive in negative u-strain while it is quite robust in positive u-strain and t-strain the FeCrAs. The maximum value of band gap at 0% tetragonal-strain is about 1.04 eV. The FeCrAs Heusler alloy preserve the 100% spin polarization for whole studied range of tetragonal-strain (−10% to 10%). This property may make this system a promising candidate for spin transfer torque application.
© 2021 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the 1st International Conference on Computations in Materials and Applied Engineering – 2021.

1. Introduction

Currently multiple fields/technologies are emerging which have the potential of replacing electron-based devices such as; Spintronics [1], Photonics [2] & Polaritons [3]. Spintronic devices, based on spin-assisted electronics, have desirable characteristic such as less power consumption and high operational speed along with highly dense material which makes it more prominent for next generation semiconductor technology [4–6]. High spin polarization is the prerequisite of spintronics devices and half-metals [7] are a class of materials which can meet this precondition. Half-metals have 100% spin polarization due to their exotic bandstructure i.e., one spin channel shows metallic nature and the other spin channel has a band gap at Fermi level. Heusler alloys are currently at the centre of extensive research because these materials possess robust half-metallic ferromagnetism [4,8–11].

There are four type of stoichiometric Heusler alloys currently existing in literature, namely, full-Heusler alloy, semi Heusler alloy, inverse Heusler alloy, and quaternary Heusler alloy. Half-metallic ferromagnetism was first predicted, in 1983, by de Groot et al. in NiMnSb semi Heusler alloy [12] and since then this property has been extensively studied in various Heulser alloys both theoretically and experimentally [13–21]. Semi-Heusler alloys are specifically important because of their high Curie temperature which makes them potential candidates for spintronics devices such as giant-magneto resistance (GMR), tunnel-magneto resistance (TMR) and magnetic tunnel junction (MTJ) [22–25].

The half-metallic ferromagnetism is a very delicate property which often attributes to the stochiometric structures in Heusler alloys. However, for application of spintronics devices such as MTJ, the thin films of half-metallic Heusler alloys are needed, therefore the half-metallic ferromagnetism must sustain in thin film structures. It is shown by Sozinov et al. [26] that induced strain in thin film structures. This strains from adjacent layers are very significant and may lead to non-uniform strains in the Heusler thin films. This induced strain may lead to the potential lattice mismatch between the semiconductor substrate and the thin films leading to tetragonal distortion of the thin-films which most likely to affect the half-metallicity of the system [27]. Therefore, the perseverance of half-metallic ferromagnetism under the effect of strains in these materials should be studied. Keeping this and the importance of semi-Heusler alloys in mind, we have chosen a recently proposed semi-Heusler alloy (FeCrAs) [28] to

* Corresponding author.
E-mail address: mukhtiyarsingh@dtu.ac.in (M. Singh).

https://doi.org/10.1016/j.matpr.2021.05.572
2214-7853/© 2021 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the 1st International Conference on Computations in Materials and Applied Engineering – 2021.

Please cite this article as: A. Singh, M. Singh and M. Singh, A *first principles* based study of the effect of uniform and tetragonal strains on half-metallicity in FeCrAs Heusler alloy, Materials Today: Proceedings, https://doi.org/10.1016/j.matpr.2021.05.572

analyze the effect of strains on half-metallic ferromagnetism of this system. FeCrAs Heusler alloy is first time used to study the effect of uniaxial and tetragonal strain to best of our knowledge.

2. Computational details

The effect of strain on electronic structures of FeCrAs semi-Heusler alloy has been carried out using density functional theory based *ab-initio* method as implmented in Wien2k program package [29]. It is a full-potential method which is one of the most accurate electronic structure calcultion methods. The generalized gradient approximation (GGA) formalism within Perdew-Burke-Ernzerhof parameterization [30] has been used to deal with exchange and correlation potentials. The plane wave cut-off parameters have been decided by $R_{MT}k_{max} = 8$ and $G_{max} = 12$ a.u.$^{-1}$ has been used for Fourier expansion of potential in the interstitial region. The energy convergence criterion of 10^{-6} Ry has been used to get exact ground state and charge convergence was also monitered along with energy. A uniform k-mesh grid of $17 \times 17 \times 17$ has been used in the Irreducible Brillouin Zone (IBZ). We have applied two types of strains i.e. uniform (u) and tetragonal (t) strain. These strains have been simulted by keeping the volume of the cell constant.

3. Results and discussion

3.1. Structural optimization and electronic properties

FeCrAs semi-Heusler alloy crystallizes in $C1_b$-structure with space group no. 216, Fe sits at (0, 0, 0); Cr at (0.25, 0.25, 0.25) and Z at (0.75, 0.75, 0.75) with vacant position at (0.5, 0.5, 0.5). The optimized equilibrium lattice parameter is 5.49 Å which is in good agreement with the previous study [28]. All further calculations are performed on this lattice parameter. The band-structure and total density of states (TDOS) of unstrained FeCrAs is shown in Fig. 1. It shows that the spin up channel is completely occupied at Fermi-level, which represents metallic behaviour. Whereas the spin down channel is completely vacant at Fermi-level, showing characteristic of semi-conducting behaviour. In spin-down, the valence band minimum is lying at −0.71 eV which is along high symmetry point 'L' and conduction band maximum is at 0.33 eV which is along high symmetry point 'X'. This results in the minority band gap of 1.04 eV which is of indirect nature. The calculated value of minority band gap is in good agreement with reports available in the literature.

3.2. Effect of strain on the electronic properties

After analysing the electronic structure of unstrained FeCrAs, next we discuss the effect of u- and t-strain. In Fig. 2, the variation of total energy versus concertation of strains is presented. For both strains, an almost uniform change is observed in energy with decrease and increase of strain concentration, except a sudden dip in energy at −8% of uniform strain. This dip may be attributed to the fact that around this strain value an overall antiferromagnetic ordering is stable in the system.

To discuss the sustainability half-metallicity under strain, we represent profile of band-gap with strain in Fig. 3. For −10% to −4% of u-strain, a complete loss of half-metallicity is observed as the minority band gap is zero in this range (Fig. 3). But from −2% to 10% a sizable band gap is observed which leads to 100% spin polarization for this range of u-strain. The maximum band gap of 1.20 eV is observed at 4% u-strain. For t-strain (Fig. 4) there is an increment in band gap exhibited from −10% to 0% and decrement in band gap exhibited from 0% to 10%. The maximum value of band-gap in t-strain is about 1.04ev at 0%. It is important to emphasize here that for whole range of t-strain (−10% to 10%), FeCrAs Heusler alloy gives 100% spin polarization. This property may make this system a promising candidate for spin transfer torque application.

We also observe the shift of indirect band-gap in u- and t-strain (Figure not shown here). In u-strain, from −2% to 2%, 'L-X' indirect bandgap is observed whereas from 4% to 10% an indirect band gap along and 'Γ-X' direction is observed. In t-strain, we observe an indirect band gap along 'Γ-N' direction for −10% to 4% and for 6% to 10% its shifted along 'Γ-H' direction. The energy of conduction band and valence band pushes upward in positive u-strain, while pushing down in negative u-strain. Hence, the loss of half-metallicity is found soon in negative u-strain. However, the conduction band and valence band in positive and negative t-strain push towards the Fermi-level, thus the profile (Fig. 4) of t-strain shows the almost symmetric curve about zero strain.

The profile of spin polarization is given in Fig. 4. It is clear that at −10%, −6% and −4% of u-strain less than 10% spin polarization can be obtained. Whereas the spin polarization of around 97% is observed at −8% of u-strain. The 100% spin polarization occurred at −2% to 10% and −10 to 10% in u-and t-strain, respectively. We can conclude from this that FeCrAs maintains the half-metallicity in t-strain and in positive u-strain, from which we can infer that the range of half-metallicity in strain is spread across quite well,

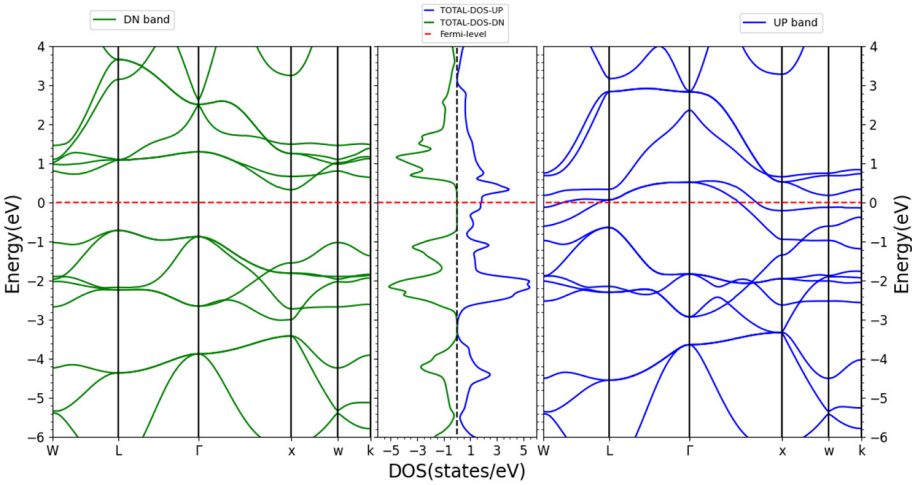


Fig. 1. Band-structure and total density of states of unstrained FeCrAs.

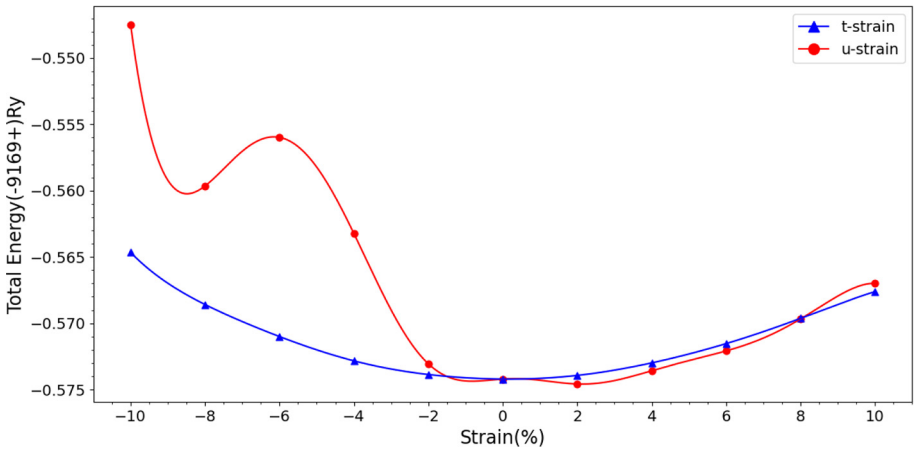


Fig. 2. Total energy versus strain (u- and t-strain).

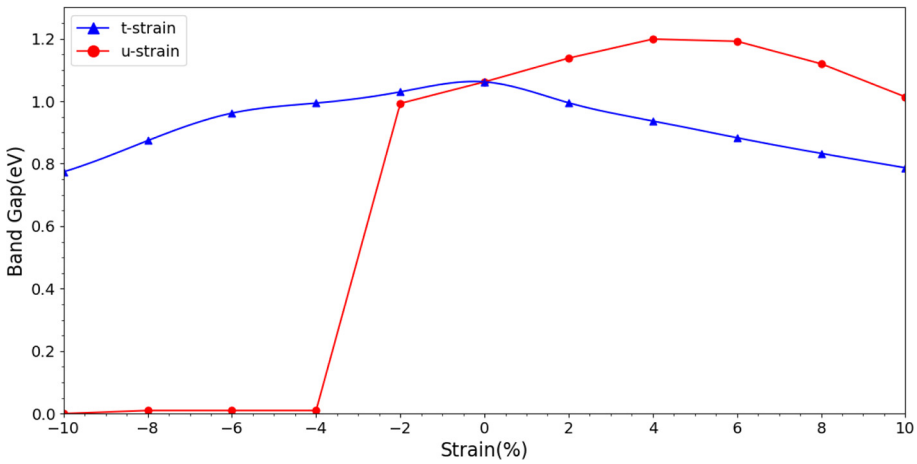


Fig. 3. Bandgap versus strain (u- and t-strain).

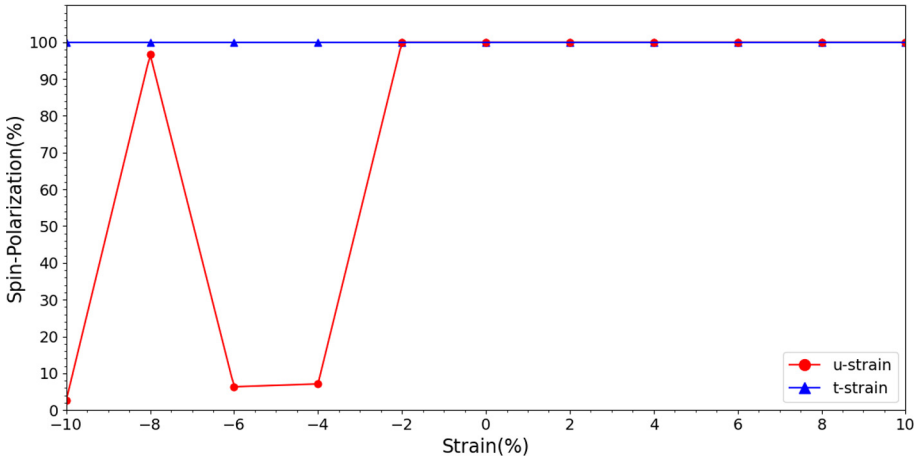


Fig. 4. Spin Polarization versus Strain (u- and t-strain).

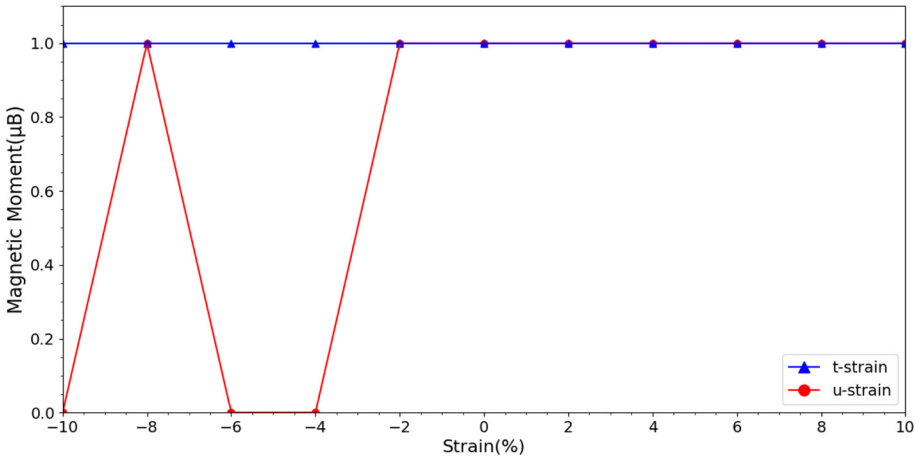


Fig. 5. Magnetic Moment versus Strain (u- and t-strain).

showing that robustness of half-metallicity is high in presence of t-strain and u-strain.

3.3. Magnetic properties

In this section, we discuss the total magnetic moment of FeCrAs under the influence of u- and t- strain. The total magnetic moment versus strain is represented in Fig. 5. The total magnetic moment of 1 μB being observed at equilibrium lattice constant of FeCrAs. It is observed that the application of u-strain does not lead to change in total magnetic moment (1 μB) in the whole range of strain except for −10%, −6% and −4% of u-strain where total magnetic moment appears to be zero. The probable cause for the same is alignment of magnetic moment in adjacent atom being opposite and approximately equal.

Although total magnetic moment of 1 μB is maintained at −8% u-strain from above trend, physical nature of FeCrAs at this particular value is metallic ferromagnetic whereas at −10%, −6% and −4% u-strain the material shows the metallic anti-ferromagnetic nature. Physical nature of FeCrAs at −2% to −10% is half-metallic ferromagnetic. Application of t- strain does not lead to change in total magnetic moment from −10% to 10% strain. It remains same as the value at equilibrium strain, which depict that t-strain preserve the half-metallic ferromagnetic nature of FeCrAs for whole studied range i.e., −10% to 10%.

4. Conclusion

We have presented here a systematic analysis of effect of uniform and tetragonal strains on electronic and magnetic properties of FeCrAs Heusler alloy. Previous works focusing on study of strain effect have not used the FeCrAs system. We have found that the system under investigation showed that its half-metallic nature has a mixed response to both of the strains. The system is also found to display the change in magnetic properties in u-strain, exhibiting the AFM and FM ordering in u-strain while only showing FM characteristic in t-strain. The Half-metallicity of FeCrAs in t-strain is maintained over a range of −10% to 10%, showing the robustness of FeCrAs in t-strain. The half-metallicity of FeCrAs is exhibited all over in the positive u-strain while in negative u-strain, the loss of half-metallicity is observed, showing the sensitivity of FeCrAs in negative u-strain and consistency in positive u-strain. In −10% to −4% of u-strain, less than 5% spin polarization has observed, except at −8%, where 97% spin polarization has exhibited. These results help us conclude that FeCrAs Heusler alloys could be a good material for spintronics application because it retains its half-metallic character under wide range of tetragonal strain.

CRediT authorship contribution statement

Aniket Singh: Investigation, Data curation, Writing - original draft, Software. **Mansher Singh:** Data curation, Writing - original draft, Validation. **Mukhtiyar Singh:** Supervision, Conceptualization, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

All the authors are highly thankful to Delhi Technological University, Delhi, India for providing the computational resources for this work.

References

[1] A. Hirohataa, K. Yamadab, Y. Nakatanic, I.L. Prejbeanud, J. Magn. Magn. Mater. 509 (2020) 16671.
[2] J. O'Brien, J. Sci. AAAS 318 (2007) 1567.
[3] M. Ram, A. S., Abeer E. Aly, A. Shankar, J. RSC. 10(2020) 7661-7670.
[4] L. Tonks, I. Langmuir, Phys. Rev. 33 (1929) 195–210.
[5] T. Kobayashi, J. Salfi, C. Chua, J. van der Heijden, M.G. House, Nat. Mater. 20 (2021) 38–42.
[6] I. Žutić, J. Fabian, S. Das Sarma, Rev. Mod. Phys. 76 (2004) 323.
[7] J.J. Attema. Chioncel C.M. Fang G.A. de Wijs, R.A. de Groot, Half-Metals: Challenges in Spintronics and Routes Toward Solutions. In: Donath M., Nolting W. (eds) Local-Moment Ferromagnets. Lecture Notes in Physics, vol 678. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11417255_13.
[8] A. Kundu, S. Ghosh, R. Banerjee, S. Ghosh, B. Sanyal, Sci. Rep. 1803 (2017) 2074.
[9] M. Singh, H.S. Saini, M.K. Kashyap, J. Alloy. Comp. 580 (2013) 201–204.
[10] L. Hongzhi, Z. Zhiyong, M. Li, X. Shifeng, L. Heyan, Q. Jingping, W. Guangheng, J. Phys. D: Appl. Phys. 40 (2007) 7121.
[11] X.J. Zhang, Z.H. Liu, Y.J. Zhang, H.Y. Liu, G.D. Liu, Y.T. Cui, X.Q. Ma, Intermetallics 73 (2016) 26–30.
[12] R.A. de Groot, F.M. Mueller, P.G. van Engen, K.H.J. Buschow, Phys. Rev. Lett. 50 (1983) 2024.
[13] M. Zhang, X. Dai, H. Hu, G. Liu, Y. Cui, Z. Liu, J. Chen, J. Wang, G. Wu, J. Phys. Condens. Matter 15 (2003) 7891.
[14] I. Galankis, K. Ozdoğan, E. Sasioglu, J. Appl. Phys. 104 (2008) 083916.
[15] I. Galanakis, P.H. Dederichs, N. Papanikolaou, Phys. Rev. B 66 (2002) 174429.
[16] H.C. Kandpal, G.H. Fecher, C. Felser, J. Phys. D: Appl. Phys. 40 (2007) 1507.
[17] M. Singh, H.S. Saini, Manish K. Kashyap, Comp. Mater. Sci. 53 (2012) 431–435. P.J. Webster, K.R.A. Ziebeck, in: H.R.J. Wijn (Ed.), Alloys and Compounds of d-Elements with Main Group Elements. Part 2, Landolt-Bornstein, New Series, Group III, 19, Springer, Berlin, 1986, pp. 75–184.

[18] K.E.H.M. Hanssen, P.E. Mijnarends, Phys. Rev. B 34 (1990) 5009.
[19] K.E.H.M. Hanssen, P.E. Mijnarends, L.P.L.M. Rabou, K.H.J. Buschow, Phys. Rev. B 42 (1990) 1533.
[20] J. Androulakis, S. Gardelis, P.D. Buckle, J. Giapintzakis, Appl. Phys. A 79 (2004) 1211.
[21] S. Gardelis, J. Androulakis, O. Monnereau, P.D. Buckle, J. Giapintzakis, J. Phys. Conf. Ser. 10 (2005) 167.
[22] S. Datta, B. Das, Appl. Phys. Lett. 56 (1990) 665.
[23] C.T. Tanaka, J. Nowak, J.S. Moodera, J. Appl. Phys. 86 (1999) 6239.
[24] J.A. Caballero, Y.D. Park, J.R. Childress, J. Bass, W.-C. Chiang, A.C. Reilly, W.P. Pratt Jr., F. Petroff, J. Vac. Sci. Technol. A 16 (1998) 1801.
[25] C. Hordequin, J.P. Nozieres, J. Pierre, J. Magn. Magn. Mater. 183 (1998) 225.
[26] A. Sozinov, A.A. Likhachev, N. Lanska, K. Ullakko, Appl. Phys. Lett. 80 (2002) 1746.
[27] T.C. Chibueze, C.E. Ekuma, A.T. Raji, F.I. Ezema, C.M.I. Okoye, J. Alloy. Comp. 848 (2020) 156186.
[28] R. Djelti, A. Besbes, B. Bestani, J. Optic. Quant. Elect. 52 (2020) 414.
[29] P. Blaha, K. Schwarz, G.K.H. Madsen, D. Kvasnicka, J. Luitz, WIEN2k, An augmented plane wave + Local Orbitals Program for calculating Crystal Properties, Karlheinz Schwarz, Techn. Universitat Wien, Wien, Austria, 2001.
[30] J.P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. 77 (1996) 3865.

Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021)
IEEE Xplore Part Number: CFP21K74-ART; ISBN: 978-0-7381-1327-2

A Literature Review on Brain Tumor Detection and Segmentation

Aditya Miglani
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
adityamiglani@outlook.com

Hrithik Madan
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
madanhrithik@gmail.com

Saurabh Kumar
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
saurabh_dtu@outlook.com

Sanjay Kumar
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
sanjay.kumar@dtu.ac.in

2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) | 978-1-6654-1272-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICCS51141.2021.9432342

***Abstract*—A tumor is a swelling or abnormal growth resulting from the division of cells in an uncontrolled and disorderly manner. Brain tumors are an exceptionally threatening kind of tumor. There exist several types of brain tumors which are classified into four grades. The process for the medical treatment of brain tumors depends on the type, the grade as well as the location of the tumor. If not detected at the early stages, brain tumors can turn out to be fatal. Magnetic Resonance Imaging (MRI) images are used by specialists and neurosurgeons for the diagnosis of brain tumors. The accuracy depends on the experience and domain knowledge of these experts, and is also a time consuming and expensive process. To overcome these restrictions, several deep learning algorithms have been proposed for the detection of presence of brain tumors. In this review paper, an extensive and exhaustive guide to the sub-field of Brain Tumor Detection, focusing primarily on its segmentation and classification, has been presented by comparing and summarizing the latest research work done in this domain. This research work has made a comparison between 28 research papers and highlighted the different state-of-the-art approaches. With a lot of ongoing research work in this area, this paper would assist all future researchers.**

***Keywords*—Brain Tumor Detection, Deep Learning, Image Segmentation, MRI images, Medical Imaging, Tumor Segmentation.**

I. INTRODUCTION

The brain is one of the most vital and sensitive organs of the body and is the centre of all nervous activity. Issues pertaining to the brain are widely considered the most difficult to operate on. There are roughly 350,000 new brain tumor cases across the world each year and the 5-year survival rate for people diagnosed with a brain tumor is only 36% [11]. Brain tumors can be classified as benign (non-cancerous tumors) or malignant (cancerous tumors) [18]. The World Health Organisation (WHO) has classified brain tumors into four grades from Grade I to Grade IV depending on the severity of the case. Neurosurgeons often recommend surgery for the treatment of brain tumors. However, for the latest stages of the tumors, alternative approaches using radiation and chemotherapy are often suggested, since the only possible diagnosis is to try to kill or slow down the growth of the cancerous cells [15]. Since the fatality rate of people diagnosed with brain tumors is very high, the detection of brain tumors at the early stages is extremely important for correct treatment. For diagnostic purposes, images can be gathered from several medical imaging techniques. Some of the medical imaging techniques are PET(Positron Emission Tomography), MRI(Magnetic Resonance Imaging), CT(Computed Tomography) [13]. Among these, MRI is considered to be the most effective medical imaging technique, especially for taking a look at soft tissues and the nervous system. MRI makes use of a powerful magnet and radio waves to produce images of body structures by causing them to emit faint signals. Unlike CT scans, MRI does not use radiation of X-rays which can be damaging. MRI images have high resolution and are quite detailed, due to which they can even detect tiny things [6, 17]. The tumor may show up as a white area, or bright white-coloured pattern. However, there are other parts of the brain which have a similar behaviour as these cells and can lead to a wrong diagnosis.

With tens of thousands of patients suffering from brain tumors each year, the use of deep learning techniques for the purpose of automatic detection and classification of brain tumors has become an area of interest. These techniques have also been employed for the segmentation of brain tumors, and this area is getting widespread attention from the medical community [16]. The aim of segmentation is to change the representations of the different areas of an image, making areas of the image having different characteristics easier to interpret. By dividing the image of the brain into these different and unique areas, each area becomes spatially contiguous [1, 2]. Some of the common problems in manual detection of the brain tumors are the significant time requirements and the possibilities of misclassifications due

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:32:05 UTC from IEEE Xplore. Restrictions apply.

to the complexity of the problem. Therefore, the automatic segmentation of MRI images of brain tumors can significantly improve the diagnosis and methods of treatment, especially in cases where access to trained experts and radiologists is difficult [3]. This paper will discuss about the latest research work done in this domain.

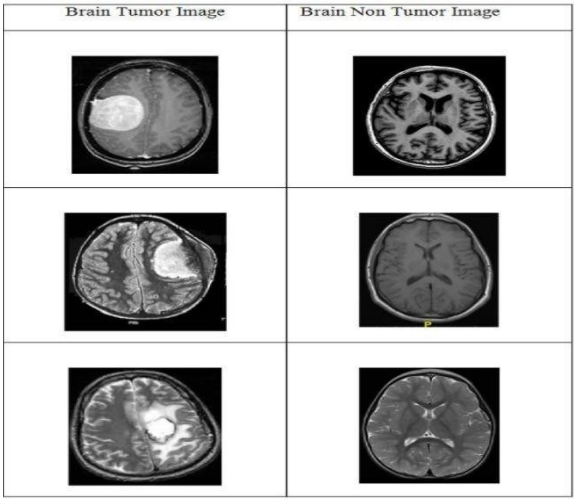


Fig. 1. Positive vs Negative Image.

II. Literature Survey

Research work by 9 different authors has been discussed on the basis of varied deep learning techniques and architectures adopted by them.

Sakshi Ahuja et al.,[1] used transfer learning and superpixel technique for detection of brain tumor and brain segmentation respectively. The dataset used was from BRATS 2019 brain tumor segmentation challenge and this model was trained on the VGG 19 transfer learning model. Using the superpixel technique the tumor was divided between LGG and HGG images. This resulted in an average of dice index of 0.934 in opposition to ground truth data.

Hajar Cherguif et al.,[2] used U-Net for the semantic segmentation of medical images. To develop a good convoluted 2D segmentation network, U-Net architecture was used. BRATS 2017 dataset was used for testing and evaluating the model proposed. The U-Net architecture proposed had 27 convolutional layers, 4 deconvolutional layers, Dice_coef of 0.81.

Chirodip Lodh Choudhury et al.,[3] made the use of deep learning techniques involving deep neural networks and also incorporated it with a Convolutional Neural Network model to get the accurate results of MRI scans. A 3-layer CNN architecture was proposed which was further connected to a fully Connected Neural Network. F-score equal to 97.33 and an accuracy equal to 96.05% was achieved.

Ahmad Habbie et al.,[4] MRI T1 weighted images were taken and using semi automatic segmentation analyzed the possibility of a brain tumour using an active contour model. The performance of morphological active contour without edge, snake active contour and morphological geodesic active contour was analyzed. MGAC performed the best among all three as suggested by the data.

Neelum et al.,[5] used a concatenation approach for the deep learning model in this paper and the possibility of having a brain tumor was analyzed. Pre trained deep learning models which are Inception - v3 and DenseNet201 were used to detect and classify brain tumors. Inception - v3 model was pre trained to extract the features and these features were concatenated for tumor classification. Then, the classification part was done by a softmax classifier.

Ms. Swati Jayade et al.,[6] used Hybrid Classifiers. The classification of tumors was done into types, malignant and benign. Feature dataset here was prepared by Gray level Co-occurrence Matrix (GLCM) feature extraction method. A hybrid method of classifiers involving KNN and SVM classifiers was proposed to increase efficiency.

Zheshu Jia et al.,[7] the author made a fully automatic heterogeneous segmentation in which SVM (Support Vector Machine) was used. For training and checking the accuracy of tumor detection in MRI images, a classification known as probabilistic neural network classification sytem had been used. Multi spectral brain dataset is used and this model focused on the automated segmentation of meningioma.

DR. Akey Sungheetha, DR. Rajesh Sharma R.[8] used Gabor transform along with the soft and hard clustering for detecting the edges in the CT and MRI images. A total of 4500 and 3000 instances of MRI images and CT were used respectively. K-means clustering was used for the separation of similar features into sub-groups To represent the images in the form of histogram properties, the author used Fuzzy c means.

Parnian Afshar et al.,[9] used a bayesian approach for the classification of brain tumor using capsule networks. To improve the results of tumor detection, capsule network instead of CNN was used as CNN can loose the important spatial information. The team proposed BayesCap framework. To test the proposed model they used a benchmarkbrain tumor dataset.

Further, 19 other research papers regarding the latest research work done in this domain have been reviewed and tabulated in Table 1. Key details such as the dataset and techniques used, year of publication, major observations and accuracy achieved have been presented to aid further research work in this domain.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:32:05 UTC from IEEE Xplore. Restrictions apply.

Table 1: The following table contains different techniques opted by researchers on Brain Tumor Detection and Segmentation.

Year, Publication	Dataset used	Technique	Observations	Accuracy
2020,[10] (Elsevier)	BRATS 2013 BRATS 2015	FCNN	• F2 FCN was used in order to operate the pixel wisebrain tumor segmentation. • A feature reuse model was introduced in order to make thereuse rate of valuable features better.	Dice: 0.86 ±0.14 PPV: 0.89 ± 0.15 Sensitivity: 0.87 ±0.16
2020,[11] (Wiley)	BraTS2015 BraTS2017 BraTS2018	3D CNN,SVM Feature Selection Architecture, FNN,KNN, Decision Trees VGG19	• The tumor was correctly identified in poorly contrasted MRI scans using this architecture. When a model which is pertained by using the extraction of the tumor images the classification accuracy of tumor type's became better from earlier.	BraTS 2015-98.32% BraTS 2017-96.97% BraTS 2018-92.67%
2020,[12] (IEEE)	ImageNet Dataset (http://www.image-net.org/download)	HCNN, CRF-RRNN	• A deep learning technique consisting of the combination HCNN and CFR in a single system was proposed for the segmentation. • The experimental results show that integration with CRF-RRNN andHCNN can increase segmentation robustness.	Precision Ratio – 96.5% Recall Ratio – 97.8%
2020,[13] (IEEE)	Kaggle Dataset Warehouse (https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection)	CNN	• Used augmentation in the dataset and expanded its size to 14 times. • ACNN was designed for the classification of MRI images. Images are labeled , label 1(having tumor) and label 0(not havingtumor).	ACNN – 96.7%
2020,[14] (Wiley)	BraTS 2015 (https://www.smir.ch/BRATS/Start2015)	CNN,MKKMC	• The performance of the proposed method is evaluated using PPV, NPV, FPR, FNR, sensitivity, specificity and accuracy. • To analyse the performance of the model the proposed method iscompared with the existing method.	Accuracy - 99% Sensitivity -93.5% Specificity –99.22%
2019,[15] (IEEE)	BraTS 2013	Two path architecture, CNN, Cascade architecture	• The model was primarily focused on the disease Glioblastomas. • A new model was built which consisted of two path and cascade architecture.	Dice Index- 0.95 Sensitivity -0.925 Precision -0.961
2019,[16] (IEEE)	3D MRI dataset from the BRATS challenge in MICCAI2016.	Random Forest Classifier	• For the classification of existence of a unique class a total of five random forest classifiers were used and all the classifiers were trained in two-class fashion. • Experimental results showed a very good improvement in segmentation for all the three tumor sections.	Whole Tumor -89.53% Tumor Core - 79.86% Active Tumor - 76.92%
2019,[17] (IEEE)	MICCAI BraTS2015	Image Processing Technique, SVM, Random Forest	• Using image processing techniques Brain tumors are segmented and shape-based features are used for feature extraction . • Extracted shape-based features are given to ML algo as SVM & random forest algorithm to find malignant and benign brain tumors.	Accuracy – 86.66%
2020,[18] (Wiley)	https://figshare.com/articles/dataset/brain_tumor_dataset/1512427	CLFAHE, DBN	• The proposed model had a total of five phases in the evaluation process. • The main aim of the paper was to first select the hidden neurons in DBN and then select the bounding limits using the new hybridisation optimization algorithm	Accuracy–92.15% Sensitivity -77.5% Specificity –9926%
2020,[19] (IEEE)	https://figshare.com/articles/dataset/brain_tumor_dataset/1512427	R-CNN. RPN	• VGG-16, a CNN architecture was used as a base layer for the proposed networks(classifier network and region proposal network) • This paper work can be extended further for calculating the percentage area of the tumor with respect to brain region of human.	Glioma Tumor -75.18% Meningioma – 89.45% Pituitary – 68.18%
2020,[20] (Elsevier)	BraTS 2017 (https://www.med.upenn.edu/sbia/brats2017/registration.html)	sparse constrained level set algorithm	• The method mentioned in thi paper found out common characteristics of brain tumor shape and was able to make asparse representation model for it. • A method energy function ground on level set was constructed.	Accuracy – 96.2%
2020,[21] (Wiley)	Flair dataset, T1 dataset, T2dataset	ISOA algorithm, DBN,SVM	• An automated system was proposed for the diagnosis of brain tumors.	Seagulls Population - 150

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:32:05 UTC from IEEE Xplore. Restrictions apply.

2019,[22] (IEEE)	1667manually annotated images	Feature extracrion algorithm, CNN-Softmax	• The classification and segmentation were done through the combination of feature extraction algorithm and CNN. • RBF classifier and DT classifier were used to find the accuracy of CNN.	Accuracy – 99.12%
2019,[23] (Elsevier)	BraTS, Sim BraTS	Dolphin-SCA algorithm	• Dolphin-SCA based depp CNN (a deep learning method) was proposed. • A total of four steps were followed including pre-processing, segmentation, feature extraction and its classification.	Accuracy – 96.2% Sensitivity –99.2% Specificity –91%
2019,[24] (ACM)	BraTS 2015 (https://www.smir.ch/BRATS/Start2015https://www.smir.ch/BRATS/Start2015)	Deep Learning, CNN	• The proposed architecture was tested on MRI images to identify the patients suffering from various type of tumors.	Dice Similarity Index – 86.7% Accuracy – 98.33%
2019,[25] (ACM)	BraTS dataset	Deep Learning, Fuzzy K-Means Clustering,	• The KMFCM was used as it can deal with higher number of segmentation problems and minimal execution time. • The model was evaluated on the basis of white pixels , black pixels, processing time, and the tumor detected area.	Accuracy – 97.3%
2020,[26] (Springer)	BraTS dataset	Deep Learning, Fuzzy K-Means Clustering, ANN	• For the segmentation of the brain tumor a model with the combination of both ANN and Fuzzy K-means algorithm was introduced. • When compared to K-Nearest Neighbour methodology the overall accuracy of the technique was increased by 8%.	Accuracy - 94% Sensitivity - 98% Specificity – 99%
2019,[27] (Springer)	BraTSdataset	FractionalPartial Differential Equation(FPDE)	• Fractional calculus approach was used to achieve the accuracy which provided an abritary order of derivative. • Mesh free technique was been used to solve the proposed equaiton to achieve higher and faster response.	Accuracy – 85%
2020,[28] (Springer)	MMRI dataset	MKSVM algorithm , K-means clustering algorithm	• In segmentation process, MRI images were preprocessed first and then the feature extraction procees was proceeded using the preprocessed images. • For the extraction of the features a raised Gabor wavelet transformation was implemented. An algorithm called as rough K means clustering was introduced for segmentaion in this paper.	Accuracy rate – 0.997

III. Brain Tumor Detection

The brain is the most sophisticated organ in the human body. It is a mass of nerve tissue composed of roughly 100 billion neurons. The brain is responsible for integrating sensory information, motor responses and is the center of learning. The Central Nervous System is made up of the brain and the spinal column. The CNS is responsible for the control of the vital functions of the body such as thought, speech and body movements. A brain tumor is an abnormal growth caused by the uncontrolled multiplication of cells. This abnormal growth in the brain or central spine can disrupt the normal functioning of the body and can affect the way a person talks, moves and processes thoughts. There are two types of brain tumors, primary tumors and secondary tumors. Primary brain tumors originate from the brain, and can further be classified as low grade or high grade. Low grade tumors grow at a slower pace as compared to high grade tumors. Secondary brain tumors are cancerous, and can start from some different part of the body and then spread to the brain [2].

The brain has a highly complex structure, making the diagnosis of brain tumors difficult

Brain tumors take the lives of nearly 250,000 people each year. Accurate diagnosis of brain tumors can save the lives of people to some extent. MRI images are extensively used to detect brain tumors. The following steps are involved in detection and treatment of brain tumors:

1. Pre-processing: This is done to improve the quality of the raw MRI images and transform them into a form, suitable for processing by humans or machines. This step also helps in removing undesired noise and enhancing overall appearance of the MRI images. Image pre-processing involves steps such as creating functions to load image datasets into arrays, resizing raw images to an established base size before feeding it to the neural network, applying normalization to rescale the pixel values so they lie within a fixed range, data augmentation to increase the size of the dataset if insufficient number of images are available, among other steps [19, 28]. These preprocessing tasks help improve classification accuracy and also speed up the training process.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:32:05 UTC from IEEE Xplore. Restrictions apply.

2. Skull Stripping: This is the process of removing parts of the images which have non-brain tissues. Through this, cerebral tissue which is not required for the analysis of the brain tumor, such as fat or skin, is removed [18]. Some popular techniques for skull stripping are based on segmentation and contouring of images.

3. Segmentation: This step aims to differentiate abnormal brain tissue from the normal brain tissue. There are manual, semi-automatic and fully automatic segmentation techniques [20]. In manual segmentation, the outline of the affected tissue area is manually traced. This method has the highest accuracy, however, it is time-consuming and cumbersome. Semi-automatic segmentation involves the users inputting some initial data to obtain the final results. In fully automatic methods, the values of the parameters do not have to be set manually and these methods can automatically detect and segment the brain tumor.

4. Feature Extraction: This step will improve accuracy of the system by selecting the more prominent features for us. It is used as a method for dimensionality reduction, and the initial data is reduced into a format which is more suitable for processing.

5. Post-processing: Post-processing provides an insight into the brain image of the tumor area. This step can include methods such as limits on the shape of the samples, contextual limitations for more accurate results and spatial control.

Brain scans can be done through different methods, the most common of which is MRI scans. The proposed algorithms discussed in this paper make the use of MRI scans, and involve the discussion of three broad tumor types:

1. **Benign Tumor:** Benign tumors are not cancerous and do not spread to other parts of the body or invade adjoining tissues. They grow at a slower pace, however, benign tumors can pose a serious problem if they press on nerves, restrict the blood flow or crowd the normal parts of the brain [17]. They generally respond well to treatment and can be removed through surgery in most cases. These tumors have low chances of recurrence.
2. **Pre-malignant Tumor:** It is not necessary that benign tumors turn cancerous, they might not. However, they can turn cancerous if the uncontrolled multiplication of tumor cells continue. Such types of tumors need to be carefully monitored for changes in the cell such as the cell appearance and growth rate.
3. **Malignant Tumor:** Malignant tumors are cancerous and can invade nearby tissues. The cancerous cells may break away from the tumor, and can then spread to other parts of the body through the lymphatic system or the bloodstream [22]. This is known as metastasis. Malignant tumors grow rapidly and can also recur, not necessarily at the same area where they initially appeared. Treatment of this type of tumors requires aggressive treatment methodologies which may include chemotherapy, radiation techniques and surgery. They are life threatening and necessitate some sort of treatment.

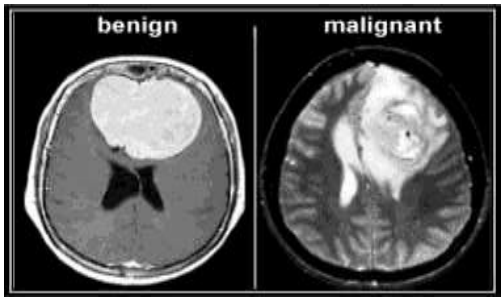


Fig. 2. Benign Tumor vs Malignant Tumor

IV. Image Segmentation

Image Segmentation involves partitioning images into its constituent segments or regions, which aids in detecting the objects and edges of the images. This partitioning is done on the basis of characteristics of the pixels in the image [16, 18]. The partitioning could separate the foreground from the background or group together pixels if they have similar shapes and colours. Image Segmentation forms an integral part of the field of medical imaging and is used extensively. Some of the common segmentation techniques are:

1. **Threshold based segmentation:** This is one of the simpler segmentation techniques. It replaces the pixels of the images with either black or white. In this technique, the comparison of the pixel value with a threshold value is done. If the value of the pixel is lower than the threshold value, the pixel is replaced with black colour, otherwise it is replaced with white. The value of the threshold can be changed as per the requirements. It is commonly used to separate the foreground and background, but the division is always into only 2 classes which is a drawback. This method can prove to be useful if the objects in question have more intensity than the background or the unwanted portions of the image.
2. **Edge based segmentation:** This technique detects edges in an image, which in turn can be used to identify certain objects [8]. Two of the common edge segmentation techniques are sobel and canny edge algorithms.
3. **Clustering based segmentation:** This technique creates segmented images from a rough initial pixel clustering. With the help of gradient ascent methods, these clusters are refined until the image is segmented. These methods try to minimize the distance between the pixels and the clusters formed [14, 28]. K- means clustering, SLIC, watershed etc. are commonly used clustering algorithms.
4. **Graph based segmentation:** In graph based segmentation, the individual pixels are considered as nodes of a graph. The degree of similarity between the adjacent pixels is then proportional to the weights of the edges between these nodes of the graph. Using the set of nodes and edges, pixels are grouped into superpixels or distinct segments. Graph cut and Normal cut are two commonly used graph based segmentation techniques.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:32:05 UTC from IEEE Xplore. Restrictions apply.

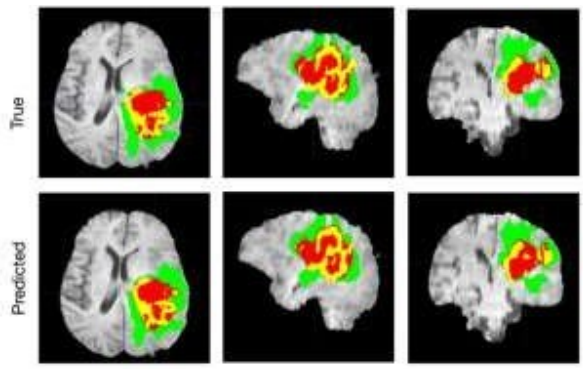


Fig. 3. Fully Automatic Brain Tumor Segmentation

V. Magnetic Resonance Imaging

Magnetic Resonance Imaging is a medical imaging technique in which radio waves generated by a computer and strong magnetic fields are used to provide detailed data and images of different parts and structures of the body. Unlike CT scans, MRI does not make use of damaging radiation of X-rays and is of no harm to the human brain [19]. It helpsin providing a cross-sectional image of the brain, which can be evaluated to ascertain the location and size of the tumor,if present. It is a non-invasive procedure. MRI machines are typically tube-shaped and surrounded by circular magnets. Thepatient is asked to lie on a table, which slides into the tube-shaped machine and examination is done. Normally, the water molecules in the body are arranged in a random manner. The working principle of MRI is that the magnetic field aligns the protons in the hydrogen atoms. These atoms are then exposed to a beam of radio waves, causing the protons to spin in a particular direction and emit faint signals which are received by the MRI machine. These signals are processed to form MRI images for further analysis [12]. Through this process, cross- sectional images of the brain can be created. These images can be used to locate the tumor and analyse its shape and size, and assess the best course of action to treat the brain tumor.

VI. Conclusion and Future Scope

This paper introduces the concept of brain tumor detection and segmentation and highlights and compares some of the key points of state-of-the-art approaches used in this domain. Some of the commonly used techniques are ML techniques like Fuzzy K-means clustering and Random Forests, as well as the use of CNN architectures is prevalent. In particular, S.Krishnakumar et al. [28] achieved the highest accuracy of 99.7% with the use of MKSVM algorithm on the MMRI dataset. High accuracy of classification of 99.12% was also achieved through the use of a combination of feature extraction algorithm and CNN [22].There are a few challenges encountered in further research work in this domain. Deep learning methods require large datasets for training purposes, and the lack of such large publicly available datasets is an obstacle. There is a scope for increase in the number of available datasets and improved access to them to aid future research work in this domain. Researchers have requested experts in this domain, such as neurologists and radiologists, to prepare structured labelling reports to aid further research. Another common issue is the presence of class imbalances in the types of tumor. This issue is commonly tackled through the use of data augmentation techniques by rotating or scaling down existing images [13]. At present, the majority of deep learning methods involve classification of tumor area, but the anatomical location of tumor region is not known to the network. Further research work in this domain can be directed towards incorporating this information in the neural network, possibly by feeding the entire image to the network. However, the size of brain tumor images are generally of high resolution and in the range of gigapixels, making the training of the network on such images unfeasible due to memory and computational power constraints.

References

1 Sakshi Ahuja, B.K Panigrahi, and Tapan Gandhi, “Transfer Learning Based Brain Tumor Detection and Segmentation using Superpixel Technique ,” in International Conference on Contemporary Computing and Applications, IEEE, 2020.
2 Hajar Cherguif, Jamal Riffi, Mohamed Adnane Mahrez, Ali Yahaouy, and Hamid Tairi, “Brain Tumor Segmentation Based on Deep Learning,” in International Conference on Intelligent Systems and Advanced Computing Sciences (ISACS), IEEE, 2019.
3 Chirodip Lodh Choudhary, Chandrakanta Mahanty, Raghvendra Kumar, and Borjo Kishore Mishra, “Brain Tumor Detection and Classification using Convolutional Neural Network and Deep Neural Network,” in International Conference on Computer Science, Engineering and Applications (ICCSEA), IEEE, 2020.
4 Ahmad Habbie Thias, Donny Danudirdjo, Abdullah Faqih Al Mubarok, Tati Erawati Rajab, and Astri Handayani, “Brain Tumor Semi-automatic Segmentation on MRI T1-weighted Images using Active Contour Models,” in International Conference on Mechatronics, Robotics and Systems Engineering (MoRSE), IEEEE, 2019.
5 Neelum Noreen, Sellapan Palaniappam, Abdul Qayyum, Iftikhar Ahmad, Muhammad Imran, “Attention-Guided Version of 2D UNet for Automatic Brain Tumor Segmentation,” IEEE Access, Volume: 8.
6 Ms. Swati Jayade, D.T. Ingole, and Manik D.Ingole, “MRI Brain Tumor Classification using Hybrid Classifier,” in International Conference on Innovative Trends and Advances in Engineering and Technology, IEEEE, 2019.
7 Zheshu Jia and Deyun Chen, “Brain Tumor Detection and Classification of MRI images using Deep Neural Network,” in IEEE acess, IEEE 2020.
8 Sungheetha, Akey, and Rajesh Sharma. ”GTIKF-Gabor-Transform Incorporated K-Means and Fuzzy C Means Clustering for Edge Detection in CT and MRI.” Journal of Soft Computing Paradigm (JSCP) 2, no. 02 (2020): 111-119
9 Pranian Afshar, Arash Mohammadi, and Konstantinos N.Plataniotis, “BayesCap: A Bayesian Approach to Brain Tumor Classification Using Capsule Networks,” IEEE Signal Processing Letters(Volume 27),2020.
10 Jinyan Hu, Yuan Wang, Deting Kong, Shuo Yan, Rui Zhao, Dengwang Li, Xiyu Liu, Yingchao Liu, Jie Xue, “Hypergraph membrane system based F 2 fully convolutional neural network for brain tumor segmentation,” in Applied Software Computing Journal, Elsevier, 2020.
11 Amjad Rehman, Muhmmad Attique Khan, Tanzila Saba, Zahid Mehmood, Usman Tariq, and Noor Ayesha, “Microscopic brain tumor detection and classification using 3D CNN and feature selection architecture,” in Microsc Res Tech. Wiley Periodicals LLC, 2020.
12 Wu Deng, Qinke Shi, Miye Wang, Bing Zheng, and Ning Ning, “Deep Learning Based HCNN and CRF-RRNN Model for Brain Tumor Segmentation,” in 10.1109/ACCESS.2017, IEEE, 2020.
13 Suhib Irsheidat, and Rehab Duwairi, “Brain Tumor Detection using Artificial Convolutional Neural Networks,” in 11th International Conference on Information and Communication Systems(ICICS), IEEE, 2020.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:32:05 UTC from IEEE Xplore. Restrictions apply.

14 Balakumaresan Ragupathy, and Manivannan Karunakaran, "A deep learning model integrating convolution neural network and multiple kernel K means clustering for segmenting brain tumor in magnetic resonance images,"in Int J Imaging Syst Technol. Wiley Periodicals LLC, 2020.
15 Ekam Singh Chahal, Ankur Haritosh, Ayush Gupta, Kanav Gupta and Adwitya Sinha, "Deep Learning Model for Brain Tumor Segmentation and Analysis," in 3rd International Conference on Recent Developments in Control, Automation and Power Engineering (RDCAPE), IEEE Access, 2019.
16 Moumen T.Elmegy, and Khaled M.Abo-El Magad, "A Multiple Classifiers System for AutomaticMultimodal Brain Tumor Segmentation," in IEEE Access, 2019.
17 Bhagyashri H.Asodekar, Sonal A.Gore, and A.D Thakare, "Brain Tumor Detectionbased on Shape Features of MRI using Machine Learning,", IEEE, 2019.
18 Bojaraj Leena, and Annamalai Jayanthi, "Brain tumor segmentation and classification via adaptive CLFAHE with hybrid classification ," in Int J Imaging Syst Technol. Wiley,2020.
19 Yakub Bhanothu, Anandhanarayanan Kamalakannan, and Govindaraj Rajamanickam, "Detection and Classification of Brain Tumor in MRI Images using Deep Convolutional Network," in 6th International Conference on Advanced Computing and Communication Systems (ICACCS), IEEE, 2020
20 Xiaoliang Lei, Xiaosheng Yu, Jianning Chi, Ying Wang, Jingsi Zhang, Chengdong Wu, "Brain tumor segmentation in MR images using a sparse constrained level set algorithm", in Expert Systems With Applications, Elsevier, 2020.
21 An Hu, Navid Razmjooy, "Brain tumor diagnosis based on metaheuristics and deep learning", in Int J Imaging Syst Technol, 2020 Wiley Periodicals LLC.
22 Masoumeh Siar, and Mohammad Teshnehlab, "Brain Tumor Detection Using Deep Neural Network and Machine Learning Algorithm," in 9th International Conference on Computer and Knowledge Engineering (ICCKE), IEEE, 2019.
23 Sharan Kumar "Optimization driven Deep Convolution Neural Network for brain tumor classification," in Biocybernetics and Biomedical Engineering Journal, Elsevier, 2019.
24 Abd El Kader Isselmou, Guizhi Xu, Shuai Zhang, Sani Saminu, and Imran Javaid, "Deep Learning Algorithm for Brain Tumor Detection and Analysis Using MR Brain Images ," in Proceedings of the 2019 International Conference on Intelligent Medicine and Health, ACM, 2019.
25 P.G Rajan, and C. Sundar, "Brain Tumor Detection and Segmentation by Intensity Adjustment," in Journal of Medical Systems, ACM, 2019.
26 R. Pitchai, P. Supraja, A. Helen Victoria and M. Madhavi, "Brain Tumor Segmentation Using Deep Learning and Fuzzy K-Means Clustering for Magnetic Resonance", in Neural Processing Letters, Springer 2020.
27 Saroj Kumar Chandra, Manish Kumar Bajpai, "Brain tumor detection and segmentation using mesh-free super-diffusive model", in Multimedia Tools and Applications, Springer 2019.
28 S. Krishnakumar, K. Manivannan, "Efective segmentation and classifcation of brain tumor using rough K means algorithm and multi kernel SVM in MR images", in Journal of Ambient Intelligence and Humanized Computing, Springer 2020.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:32:05 UTC from IEEE Xplore. Restrictions apply.

2021 2nd International Conference for Emerging Technology (INCET)
Belgaum, India. May 21-23, 2021

A Novel 8.5 GHz Parallel Coupled, Half Wavelength Bandpass Filter with Dumbbell-Shaped Defected Ground Structure

Aditya Agrawal*
Department of Electronics and Communications
Delhi Technological University
New Delhi, India
aditya.aggie1999@gmail.com

Adya Agarwal*
Department of Electronics and Communications
Delhi Technological University
New Delhi, India
adyaagarwal4@gmail.com

Ojas Aryam*
Department of Electronics and Communications
Delhi Technological University
New Delhi, India
ojasaryam@gmail.com

Asok De*
Department of Electronics and Communications
Delhi Technological University
New Delhi, India
asok.de@dtu.ac.in
**All authors have contributed equally*

2021 2nd International Conference for Emerging Technology (INCET) | 978-1-7281-7029-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/INCET51464.2021.9456441

***Abstract*—A novel 8.5 GHz parallel coupled bandpass filter is presented with dumbbell shaped defected ground structure which is placed in the ground plane. The circuit structure is simulated using ADS. The microstrip filter structure is analyzed by gradient variation of the line that passes through the midpoints of the conducting strips, which is varied linearly with different values of slope. The effects on the bandwidth and cut-off frequency are observed to understand the desired gradient for different uses of the linear bandpass filter. The measured and analyzed results are in good agreement with the theoretical results. The proposed structure has compact size, good bandwidth, and high cut off frequency.**

Keywords— Bandpass Filter, Parallel-Coupled, Microstrip, Defected Ground Structure (DGS)

I. INTRODUCTION

With the shift to 5G networks, it has made high frequency filters a necessity. Recently the analysis of microstrips filters with defected ground structure is popular because of its efficient design and compact size. DGS in some of the frequencies can provide a rejection band because there is an increase in the inductance for the microstrip line. In this paper a bandpass microstrip filter is proposed with a fractional bandwidth of 0.15 and cut off frequency of 8.5 GHz. In this paper various configurations are applied for the design by changing the gradient of the slope that cuts through centre points of the parallel coupled filters. This helps in obtaining appropriate gradient for the filter design according to different usability with minimal changes to the filter design.

A. Microstrip Filter

It is a conducting strip with width and length which is placed on top of a substrate with some dielectric constant ε_r, width w and thickness t. In microstrip lines, the transverse components are mostly considered because the longitudinal components are small and therefore, negligible. Microstrips lines behave similar to the transverse transmission lines and therefore, quasi-TEM approximations can be applied to the microstrip lines.

The dielectric substrate which is homogeneous replaces the air which has inhomogeneous dielectric. There are different kinds of losses in microstrips such as radiation losses, conductor losses and dielectric losses. However, the magnetic losses play a role for substrates which are magnetic such as ferrites.

The metallic enclosures are required for most microstrip circuits. The conducting side and top walls are present which affects the effective dielectric constant and the characteristic impedances. These are some of the few effects that must be considered while designing microstrip filters.

B. Microstrip Discontinuities

There are many discontinuities that are encountered in microstrips such as bends, steps, gaps, open-ends and junctions. The effect of these discontinuities should be taken into consideration and modelled accurately. Microstrips with right-angle bends can be made using T-network which is equivalent to capacitance and inductance. Gaps in microstrips are shown using equivalent series (C_g) and shunt capacitance (C_p). For open ends, the field does not abruptly stop but is slightly extended further because of the effect of the fringing field. This is modelled with a shunt capacitance (C_p) or can be modelled with a transmission line of equivalent length l.

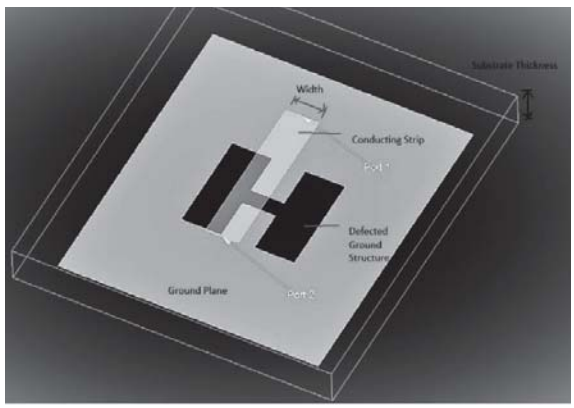


Fig. 1. General Microstrip Structure with Defected Ground Structure

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:52:09 UTC from IEEE Xplore. Restrictions apply.

C. *Parallel-Coupled, Half Wavelength Resonator Filters*

Microstrip bandpass filter with parallel coupled conducting strips has been presented in this paper. Adjacent resonators are placed such that they are along half of the length and parallel to each other. This kind of filter is convenient for building filters with a wide bandwidth as they give large coupling for a space between the resonators [1]. In [2], the suppression of 1st spurious harmonic in parallel-coupled bandpass filter with DGS has been studied.

II. Equivalent Circuit and Layout of Dumbbell DGS

Various close-packed and high-efficiency components were described [3] using structures called the defected ground structures where a defect is engraved into the ground plane of the microstrip. There are many applications of DGS in microstrip filters with various configurations [4-11]. Dumb-bell DGS is formed with 2 rectangle-shaped defects that are positioned within the ground plane and connected through a perpendicular channel that is carved in the ground. The DGS was placed according to the position and structure of our parallel coupled bandpass filter so that it doesn't overlap with the adjacent resonator which can cause a failure in the operation of the filter.

The layout of the Defected Ground Structure used is shown below in Fig. 2. There is a microstrip line above and a dumbbell shaped pattern that is carved into the ground plane.

In Fig. 2, the width of the microstrip line is 0.57 mm (w) and the characteristic impedance is 50 Ω (Z_0). The substrate used for the simulation is Rogers RT/Duroid 6010. It has substrate height h=0.635 mm, a dielectric loss tangent of 0.0023, and dielectric constant ε_r=10.2. The microstrip line is made of copper.

The dimensions of the dumbbell DGS that is used for reduction of size of the microstrip bandpass filter are a = 1.69 mm, b = 0.69 mm, and g = 0.285 mm. Simulation is done using Keysight ADS software and the simulation results are illustrated in Fig. 3.

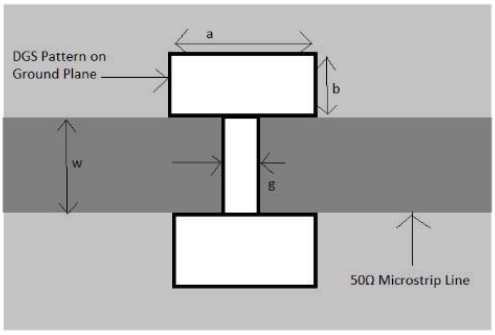


Fig. 2. Configuration of the dumbbell-shaped DGS unit.

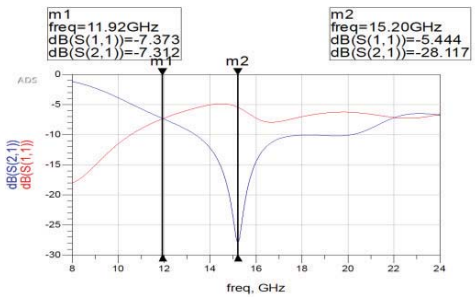


Fig. 3. Simulation result for dumbbell-shaped DGS with Microstrip Line.

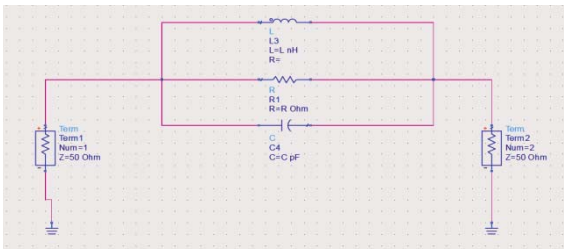


Fig. 4. Equivalent Circuit of dumbbell shaped DGS

A single parallel RLC resonator circuit (shown in Fig. 4) can be used to model the frequency characteristics of the proposed dumbbell DGS structure, where L signifies the effect of rectangle shape in the DGS unit and C signifies the effect of the channel between two rectangular shapes [8]. Changes in equivalent-circuit parameters due to variation of physical dimensions of the DGS are elaborated further in [12,13].

By using the modelling techniques in [6], the parameters of the equivalent circuit are found using (1), (2) & (3).

$$C = \frac{\omega_c}{2Z_0(\omega_0^2 - \omega_c^2)} \tag{1}$$

$$L = \frac{1}{4\pi^2 f_0^2 C} \tag{2}$$

$$R = \frac{2Z_0}{\sqrt{\frac{1}{|S_{11}(\omega_0)|^2} - \left(2Z_0\left(\omega_0 C - \frac{1}{\omega_0 L}\right)\right)^2} - 1} \tag{3}$$

Here Z_0 is the characteristic impedance of the microstrip line, f_c is the 3dB cut-off frequency, and f_0 is the resonant frequency. Values of cut-off frequency and resonant frequencies extracted from Fig. 3 are shown in Table I and the values of the parameters of the equivalent circuit, calculated using (1), (2) & (3) are also shown in Table I.

The simulation result of the equivalent circuit of dumbbell DGS (Fig. 4) is presented in Fig. 5. This result is similar to the simulation result obtained with DGS on microstrip line (Fig. 3).

TABLE I. Extracted Equivalent Circuit Parameters for Dumbbell DGS Unit

Cut-off Frequency, ω_c (GHz)	11.92
Resonant Frequency, ω_0 (GHz)	15.20
$\lvert S_{11}(\omega_0)\rvert$	0.534
L(nH)	0.514
C(pF)	0.213
R(Ω)	360

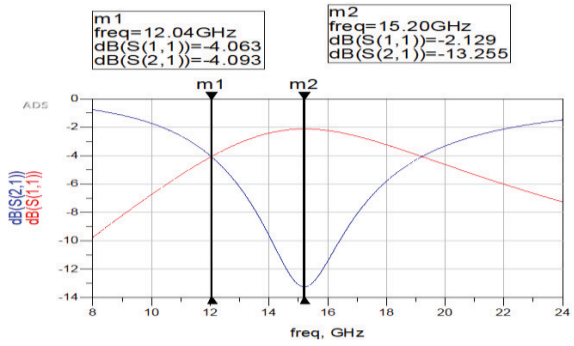


Fig. 5. Simulation Result of the equivalent RLC circuit of dumbbell shaped DGS

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:52:09 UTC from IEEE Xplore. Restrictions apply.

III. Conventional Parallel-Coupled Line Bandpass Microstrip Filter

This section is devoted to explaining the designing procedure of the parallel-coupled microstrip bandpass filter. The filter is designed with the centre frequency (f_0) chosen as 8.5 GHz and the fractional bandwidth (FBW) around f_0 as 0.15, which gives cut-off frequencies f_1=7.88 GHz and f_2=9.16 GHz for the bandpass. FBW is defined using (4).

$$FBW = \frac{\omega_2 - \omega_1}{\omega_0} \tag{4}$$

Filters can be designed using a variety of materials and by various methods [14]. The filter is chosen to be a five-pole Chebyshev Response bandpass filter with a 0.1-dB ripple in the passband. The characteristic impedance, Z_0=50 Ω. The substrate on which the filter is implemented is Rogers RT/Duroid 6010. It has substrate height h=0.635 mm, a dielectric loss tangent of 0.0023, and dielectric constant ε_r=10.2.

Element values for the low-pass prototype filter (Fig. 6) of a 5-pole Chebyshev filter with 0.1-dB ripple are g_0=g_6=1.0, g_1=g_5=1.1468, g_2=g_4=1.3712, g_3=1.975 [15].

Using these g-values, the low-pass prototype filter is converted to a band-pass filter. The 'g' elements transform to a series/parallel LC resonant circuit for given low-pass prototype circuit. For 'g' representing inductance, the elements for series LC resonator in the bandpass filter are found using (5).

$$L_s = \left(\frac{\Omega_c}{FBW\omega_0}\right)\gamma_0 g \tag{5a}$$

$$C_s = \frac{1}{(\omega_0^2 L_s)} \tag{5b}$$

For 'g' representing the capacitance, the elements for parallel LC resonator in the bandpass filter are found using (6).

$$C_p = \left(\frac{\Omega_c}{FBW\omega_0}\right)\frac{g}{\gamma_0} \tag{6a}$$

$$L_p = \frac{1}{(\omega_0^2 C_p)} \tag{6b}$$

The cut-off frequency for low-pass prototype Ω_c is chosen as 1 rad/s and impedance scaling factor γ_0=Z_0/g_0.

The RLC Bandpass Chebyshev Filter derived is shown in Fig. 7 and the values of the parameters are given in Table II.

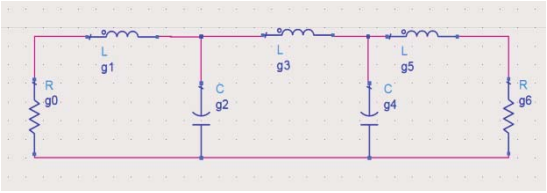


Fig. 6. Five-pole Low-pass prototype

TABLE II. Parameters of Chebyshev Bandpass Filter

i	L_i (nH)	C_i (pF)
1 and 5	7.1576	0.0489
2 and 4	0.1024	3.4233
3	12.3267	0.02844

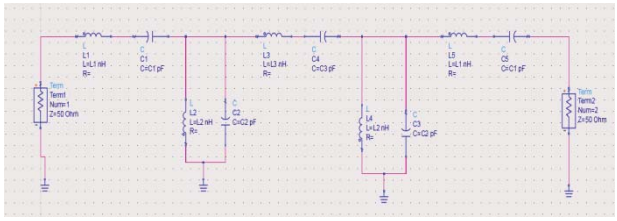


Fig. 7. Chebyshev Bandpass Filter

To design the parallel-coupled microstrip bandpass filters with half-wavelength line resonators, first the characteristic admittances of the J-inverters ($J_{j,j+1}$) are calculated using the following equations [16]:

$$\frac{J_{01}}{Y_0} = \sqrt{\frac{\pi\, FBW}{2\, g_0 g_1}} \tag{7}$$

$$\frac{J_{j,j+1}}{Y_0} = \frac{\pi FBW}{2}\frac{1}{\sqrt{g_j g_{j+1}}} \qquad j=1 \text{ to } n-1 \tag{8}$$

$$\frac{J_{n,n+1}}{Y_0} = \sqrt{\frac{\pi}{2}\frac{FBW}{g_n g_{n+1}}} \tag{9}$$

Where, Y_0= (1/50) ℧ is the characteristic admittance of the terminating line, g_i are the values of the elements of the low-pass prototype filter, and FBW is fractional bandwidth. Next, the even- and odd-mode characteristic impedances of the coupled microstrip line resonators are calculated using the following equations [16]:

$$(Z_{0e})_{j,j+1} = \frac{1}{Y_0}\left[1 + \frac{J_{j,j+1}}{Y_0} + \left(\frac{J_{j,j+1}}{Y_0}\right)^2\right] \quad j = 0 \text{ to } n \tag{10}$$

$$(Z_{0o})_{j,j+1} = \frac{1}{Y_0}\left[1 - \frac{J_{j,j+1}}{Y_0} + \left(\frac{J_{j,j+1}}{Y_0}\right)^2\right] \quad j = 0 \text{ to } n \tag{11}$$

Where Z_{0e} signifies the even mode impedance and Z_{0o} signifies the odd mode impedance. The calculated values of the impedances are shown in Table III. Using these values, the physical dimensions of the Microstrip Coupled Lines are calculated using the inbuilt calculator in ADS, LineCalc. The coupled line 1, 2 and 3 are symmetric to coupled lines 6, 5 and 4, respectively. The calculated physical dimensions for each coupled line are given in Table IV. The dimensions of microstrip line (calculated using LineCalc) are w_0=0.599322 mm & l_0=3.30971 mm.

TABLE III. Circuit Design Parameters of Parallel-Coupled Bandpass Filter

j	$J_{j,j+1}/Y_0$	$(Z_{0e})_{j,j+1}$	$(Z_{0e})_{j,j+1}$
0	0.4533	82.93666	37.60916
1	0.1879	61.16003	42.37045
2	0.1432	58.18392	43.86609

TABLE IV. Microstrip Design Dimensions of Parallel-Coupled Bandpass Filter

Coupled Line (j)	w_j (mm)	s_j (mm)	l_j (mm)
1 and 6	0.363721	0.151062	3.44776
2 and 5	0.543624	0.501387	3.36031
3 and 4	0.564447	0.666008	3.34475

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:52:09 UTC from IEEE Xplore. Restrictions apply.

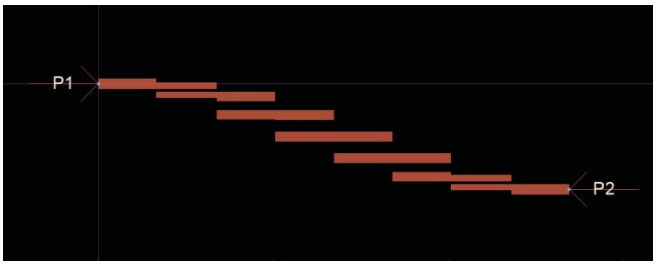


Fig. 8. Top view of the Configuration of the Parallel-coupled Microstrip Bandpass Filter

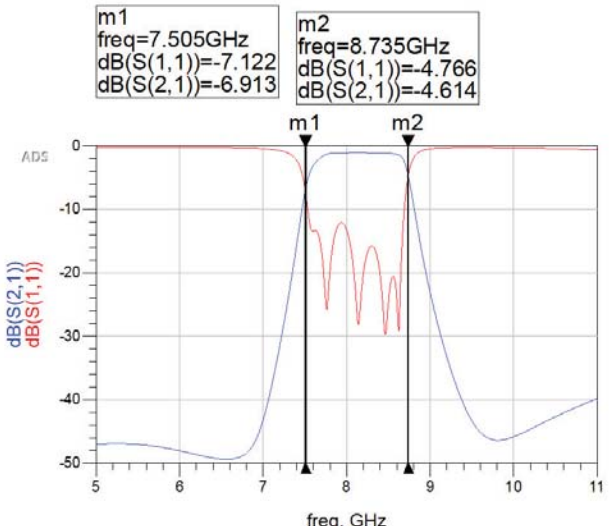


Fig. 9. Simulation of the Parallel-coupled Microstrip Bandpass Filter.

The parallel-coupled microstrip bandpass filter is designed and simulated using Keysight ADS software. The layout is given in Fig. 8 and the simulation result is given in Fig. 9. The simulation result in Fig. 9 has a centre frequency of 8.097 GHz (4.74 % error) and an FBW of 0.1519 (1.28 % error) around the centre frequency.

IV. Proposed Parallel Coupled Line Microstrip Bandpass Filter with Dumbbell DGS

For the proposed Microstrip Bandpass Filter with Defected Ground Structures (DGS), a DGS with similar specifications to the one described in Section II, is etched in the ground plane under each conducting strip at their centre points as shown in Fig. 10. All DGS sections have the same dimensions.

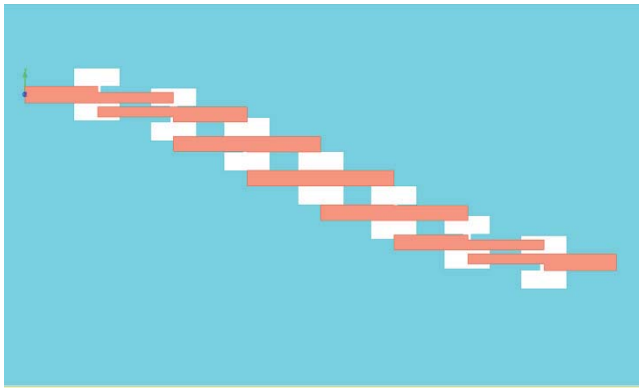


Fig. 10. Layout of the Proposed Parallel-coupled Microstrip Bandpass Filter with Dumbbell DGS.

When the Microstrip Bandpass Filter with f_0=8.5GHz has been simulated with the proposed DGS Filter Configuration (Fig. 10), the centre frequency of the simulation decreases significantly to 7.368 GHz, but the FBW remains the same around the centre frequency with a very insignificant difference.

A different parallel-coupled microstrip bandpass filter is designed & simulated using the same specifications and method as Section III but with a different centre frequency of f_0=10.5 GHz. The calculated physical dimensions for each coupled line for this new bandpass filter is given in Table V. The total length of the microstrip filter has decreased by 5.25 mm.

The microstrip parallel-coupled bandpass filter with f_0=10.5 GHz is simulated both without DGS and with DGS according to proposed design in Fig. 10. The simulation results of both are illustrated in Fig. 11 & Fig. 12.

The simulation in Fig. 11 has a centre frequency of 9.889 GHz and an FBW of 0.1522. The simulation in Fig. 12 has a centre frequency of 8.732 GHz (2.73 % error w.r.t 8.5 GHz) and an FBW of 0.1603 (6.89 % error w.r.t 0.15) around the centre frequency.

Therefore, the inclusion of DGS section in ground plane of the filter has decreased the centre frequency of the bandpass filter. The new design with DGS can simulate a bandpass filter of 8.5 GHz with smaller dimensions for the microstrip lines with the same fractional bandwidth.

TABLE V. Microstrip Dimensions of Parallel-Coupled Bandpass Filter with Centre Frequency = 10.5 GHz

Coupled Line (j)	w_j (mm)	s_j (mm)	l_j (mm)
1 and 6	0.368268	0.154844	2.77839
2 and 5	0.549329	0.517338	2.70358
3 and 4	0.569806	0.687885	2.69033

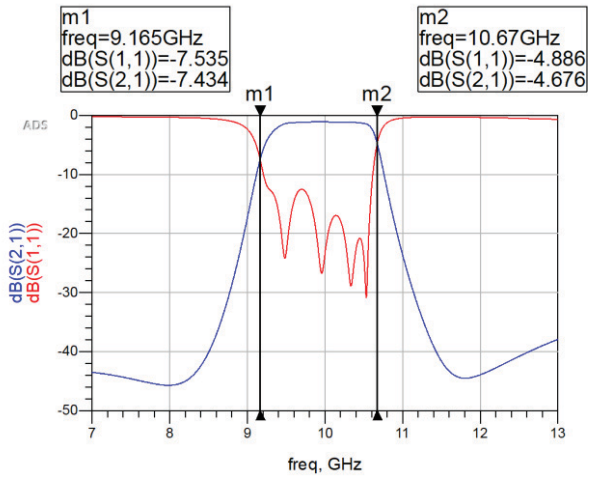


Fig. 11. Simulation of Parallel-coupled Microstrip Bandpass Filter (f_0=10.5 GHz) without Dumbbell DGS.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:52:09 UTC from IEEE Xplore. Restrictions apply.

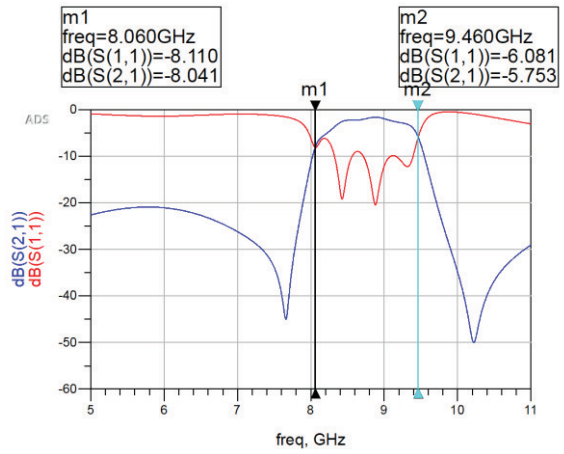


Fig. 12. Simulation of the new Microstrip Bandpass Filter with Dumbbell Defected Ground Structure.

TABLE VI. COMPARING THE RESULTS BEFORE AND AFTER INCLUDING DGS IN GROUND PLANE

Filter	Results		
	Centre Frequency (f_0) (GHz)	*FBW around f_0*	*Total Length (mm)*
f_0=8.5 GHz without DGS	8.097	0.1519	26.925 mm
f_0=10.5 GHz with DGS	8.732	0.1602	21.666 mm

V. EFFECT OF VARYING GRADIENT OF LINE THROUGH CENTRE POINTS OF MICROSTRIP LINES

In this section, the positions of the microstrip lines in the filter developed in section IV is varied along the x-axis, keeping the positions fixed w.r.t the y-axis. A straight line is drawn using the centre points of the first and the last microstrip sections, as shown in Fig. 13. The slope of this line is found to be -0.375 (absolute values are considered). All the other microstrips sections are moved so that their centres lie on this straight line. This filter is then simulated. The simulation has a centre frequency of 8.563 GHz and an FBW of 0.0794 around the centre frequency.

The slope of this straight line is varied and the positions of the microstrip sections are changed to fit the different lines, keeping the position of the top left microstrip section fixed. Two examples with slopes -0.6 & -0.25 are shown in Fig. 14 & Fig. 15. The centre frequencies and bandwidths for different slopes are shown in Table VII.

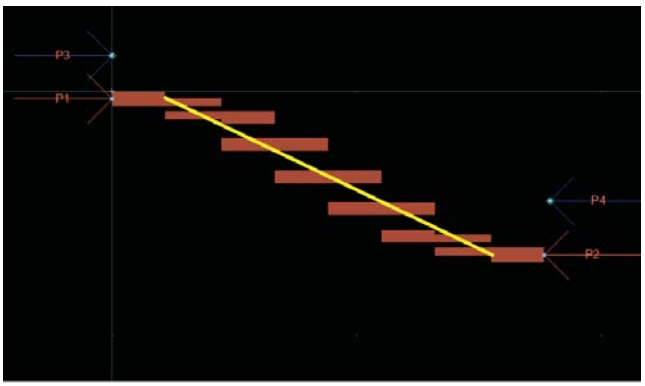


Fig. 13. Straight line drawn using the centre points of the first and the last microstrip sections.

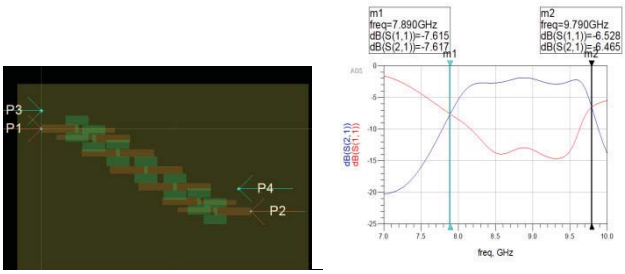


Fig. 14. Simulation Results for slope, m=-0.6

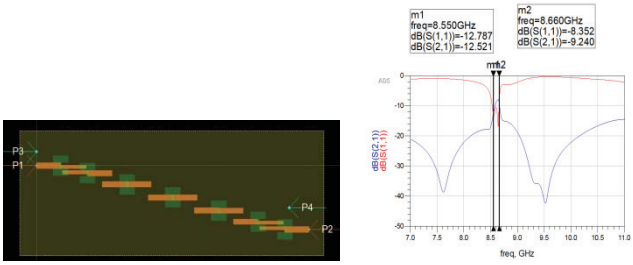


Fig. 15. Simulation Results for slope, m=-0.25

TABLE VII. RESULTS FOR DIFFERENT SLOPES OF STRAIGHT LINE

Slope (m)	Results	
	Centre Frequency (f_0) (GHz)	*FBW around f_0*
-0.6	8.7888	0.2162
-0.5	8.6718	0.1614
-0.375	8.5633	0.0794
-0.25	8.6048	0.0128

It is observed from Table VII that the bandwidth of the filter can be manipulated though variation the slope of the line that passes through the centre points of the microstrip sections. As the slope is decreased from -0.6 to -0.25, the bandwidth decreases. A very narrow bandwidth bandpass filter is obtained at slope -0.25.

VI. CONCLUSION

The study was aimed to design a filter for a centre frequency of 8.5 GHz and obtaining a minimized area for fabrication with the help of dumbbell shaped defected ground structure and after analysing data from different variations, we were able to obtain the desired result. By analysing the output waveforms for different microstrip filter design for different centre frequency, it is noticed that the increase in the centre frequency causes decrease in the area required for microstrip. To make the microstrip compact we decided to design a microstrip filter with centre frequency of 8.5 GHz. The second part of the analysis includes the variation in the gradient values for the slope that passes through the centre points of the conducting strips by keeping the upper corner constant. The bandwidth can be manipulated according to our desired output while keeping the centre frequency constant. The increase in the value of gradient causes an increase in the bandwidth but with minimal changes to the microstrip structure. Similarly, a narrow bandwidth filter can be obtained by decreasing the gradient which would result in lower bandwidth. This filter design can be further improved in the future by analysing it

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:52:09 UTC from IEEE Xplore. Restrictions apply.

with different functions along which the conducting strips are positioned.

REFERENCES

[1] S. B. Cohn. "Parallel-Coupled Transmission-Line-Resonator Filters." IRE Transactions on Microwave Theory and Techniques, vol. 6, pp. 223-231, Apr. 1958.

[2] D. Ahn, J. Park, C. Kim, J. Kim, Y. Qian, and T. Itoh, "A design of the low-pass filter using the novel microstrip defected ground structure," IEEE Trans. Microwave Theory Tech., vol. 49, pp. 86–93, 2001.

[3] L. H. Weng, Y. C. Guo, X. W. Shi, and X. Q. Chen, "An Overview on Defected Ground Structure." Progress In Electromagnetics Research, vol. 7, pp. 173-189, 2008

[4] V. Radisic, Y. Qian, R. Coccioli, and T. Itoh, "Novel 2-D photonic bandgap strucutre for microstrip lines," IEEE Microwave Guided Wave Lett., vol. 8, pp. 69-71, Feb. 1998.

[5] C. S. Kim, J. S. Park, D. Ahn, and J. B. Lim, "A novel one-dimensional periodic defected ground structure for planar circuits," IEEE Microwave Guided Wave Lett., vol. 10, pp. 131-133, Apr. 2000.

[6] D. Ahn, J. -. Park, C. -. Kim, J. Kim, Y. Qian and T. Itoh, "A design of the low-pass filter using the novel microstrip defected ground structure," in IEEE Transactions on Microwave Theory and Techniques, vol. 49, no. 1, pp. 86-93, Jan. 2001, doi: 10.1109/22.899965.

[7] R. Qiang, Wang and D. Chen, "Anovel Microstrip Bandpass Filter with Two Cascaded PBG Structures," IEEE AP-S Digest, pp. 510-513, 2001.

[8] J. P. Kim and W. S. Park, "Microstrip lowpass Filter with Multislots on Ground Plane," Electronic Letters, vol. 37, No. 25, 6 Dec. 2001.

[9] M. Kahrizi, T. Sarkar and Z. Maricevic, "Dynamic Analysis of a Microstrip line Over a Perforated Ground Plane," IEEE Trans. Microwave Theory Tech., vol. 42, No. 5, pp. 820-825 , May 1994,

[10] X. Yi, P. Yeh and J. Hong, "Photonic Bandgap Structure in Thin Film Waveguides," SBMO/IEEE MTT-S IMOC 99 Proceedings, pp. 451-453. 1999.

[11] S. Jeong, "Amplifier Design Using /4 High Impedance Bias Line with Defected Ground Structure (DGS)," IEEE MIT-S Digest, pp. 1161-1163. 2002.

[12] J. I. Park, C. S. Kim, J. S. Park, Y. Qian, D. Ahn, and T. Itoh, "Modeling of photonic bandgap and its application for the low-pass filter design," in Asia–Pacific Microwave Conf. Dig., Dec. 1999, pp. 331–334.

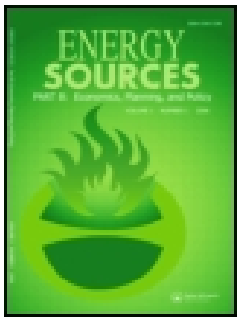
[13] D. Ahn, J. S. Park, C. S. Kim, Y. Qian, and T. Itoh, "A design of the low-pass filter using the novel microstrip defected ground structure," IEEE Trans. Microwave Theory Tech., vol. 49, pp. 86–93, Jan. 2001

[14] D. K. Mahor, "Roll-off Improvement of Open Circuit Stub Band-Stop Filter." International Journal of Research Antennas and Microwave Engineering, Vol. 3, No. 3, pp 112-118, July-Dec 2012

[15] D.M. Pozar, Microwave engineering, 2nd ed., Wiley, New York, 1998

[16] J. I. Park, C. S. Kim, J. Kim, Y. Qian, D. Ahn, and T. Itoh, "Modeling of Photonic Bandgap and Its Application for the Low-pass Filter Design." in Asia-Pasific Microwave Conference Proceedings, pp. 331-334. 1999,

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:52:09 UTC from IEEE Xplore. Restrictions apply.



Taylor & Francis
Taylor & Francis Group

Energy Sources, Part B: Economics, Planning, and Policy

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/uesb20

A novel quantum inspired hybrid metaheuristic for dispatch of power system including solar photovoltaic generation

Kapil Deo Bodha, Vinod Kumar Yadav & Vivekananda Mukherjee

To cite this article: Kapil Deo Bodha, Vinod Kumar Yadav & Vivekananda Mukherjee (2021): A novel quantum inspired hybrid metaheuristic for dispatch of power system including solar photovoltaic generation, Energy Sources, Part B: Economics, Planning, and Policy, DOI: 10.1080/15567249.2021.1933265

To link to this article: https://doi.org/10.1080/15567249.2021.1933265

Published online: 07 Jun 2021.

Submit your article to this journal

Article views: 38

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=uesb20

ENERGY SOURCES, PART B: ECONOMICS, PLANNING, AND POLICY
https://doi.org/10.1080/15567249.2021.1933265

Taylor & Francis
Taylor & Francis Group

Check for updates

A novel quantum inspired hybrid metaheuristic for dispatch of power system including solar photovoltaic generation

Kapil Deo Bodha [a], Vinod Kumar Yadav [b], and Vivekananda Mukherjee [a]

[a]Department of Electrical Engineering, Indian Institute of Technology (Indian School of Mines), Dhanbad, Jharkhand, India; [b]Department of Electrical Engineering, Delhi Technological University, Delhi, India.

ABSTRACT

This manuscript proposes a novel metaheuristic, which is an amalgamation of the quantum concept and the gravitational search particle swarm optimization technique. The inclusion of quantum concepts in the algorithm enhances its capability, as in quantum space there is no restriction on the movement of particle and the solution can be obtained with smaller population and faster convergence. An adaptive contraction expansion factor is also introduced which ensures better exploration of the algorithm. The technique is applied to solve the combined economic emission dispatch of a hybrid solar thermal unit operating in New Delhi, India for full and reduced solar radiation. To verify the effectiveness of the proposed technique, it is also applied to solve the combined economic emission dispatch of a 10 unit, 2000 MW system. It is observed that the proposed approach provides a fuel cost saving of up to 1300$ when compared with other referred techniques.

KEYWORDS

Binary encoded problems; combinatorial optimization; combined economic emission dispatch; mixed integer optimization problem; quantum evolutionary algorithm; rotation quantum gate

1. Introduction

India is the third-largest producer of thermal power in the world (Ruhe 2019) and also has one of the highest installed solar energy capacity (Bulut and Menegaki 2020). The power network in India is highly interconnected and is supplied mainly by the conventional power plants (Liu 2018). The total installed capacity in India is 370106 MW as on 31 March 2020. (Government of India,Central Eectricity Authority, Ministry of power 2020). Thermal power generation is 62.8% of total power production. Renewable energy sources (RESs) comprise of 23.5% of total installed capacity (India, Ministry of Power Govt. of 2020). During Paris summit 2015 (Nations, United 2015), India pledged to raise its renewable energy generation to 40% by the year 2032. Solar energy is available in abundance throughout the country. On an average, 58% of the country receives above 5 kWh/m^2/day of average global insolation (Ramachandra, Jain, and Krishnadas 2011). Therefore, taking into account, huge amount of solar insolation, the integration of solar energy with thermal power plant is a preferred way to reduce operating cost as well as the emission (Stavrakas, Papadelis, and Flamos 2019).

Economic load dispatch (ELD) refers to scheduling of generators in a way that they are able to supply the load at lowest fuel cost without violating any given constraint (Ravikumar et al. 2014). With increasing emphasis on clean energy for a healthier and sustainable environment, utility companies are forced to modify their design and operational strategy to reduce the emission from power plants. The operational strategy is modified by adding another objective in the conventional ELD problem. This modified ELD with the inclusion of emission is termed as combined economic emission dispatch (CEED). The CEED aims to minimize the fuel cost and the emission of the power plant without violating any constraint.

CONTACT Kapil Deo Bodha bodhas@gmail.com Indian Institute of Technology (Indian School of Mines) ,Police Line Road, Main Campus IIT (ISM), near Rani Bandh, Hirapur, Sardar Patel Nagar, Dhanbad, Jharkhand 826004

© 2021 Taylor & Francis Group, LLC

1.1. Literature review

One of the foremost techniques employed to solve CEED was the weighted min-max technique. The technique was applied on IEEE 6-generator system (Dhillon, Parti, and Kothari 1993).When real-time constraints like transmission line losses are added, the ELD problem becomes complex and difficult to solve by weighted min-max technique. To overcome this shortcoming, quadratic programming method was introduced. It was applied to solve the CEED problem (Fan and Zhang 1998). An advanced version of genetic algorithm (GA), termed as non-dominated sorting GA-II (NSGA-II), was used to find multiple Pareto optimal solutions, which is more suitable for multiobjective functions with conflicting objectives. The elitist approach consideration makes sure that the solution does not deteriorate in next generation (King and Rughooputh 2003). GA has a tendency to converge prematurely. This can be handled by using mutation strategy as proposed in (Zhang et al. 2015).

Fuel restricted economic dispatch problem is formulated using teaching-learning-based optimization (TLBO) technique (Elanchezhian, Subramanian, and Ganesan 2016). The TLBO method is based on the effect of influence of a teacher on the output of learners in a class. It was applied to solve the CEED problem of different test system considering cubic functions. The obtained results proved that the TLBO technique efficiently handles the problem and gives better results than referred counterparts. Squirrel search algorithm (SSA) is a swarm intelligence-based optimization technique. It is based on foraging behavior of squirrels for searching food. The SSA technique is utilized to solve the ELD problem including valve point effect and multiple fuel options (Sakthivel, Suman, and Sathya 2020). This technique is applied on 6-, 10-, 13-, 40-, 140-, and 160-unit test systems. The SSA offered better-quality results and exhibited excellent convergence characteristics. A Levenberg Marquardt neural network technique was applied to solve the CEED problem of a three-unit system (Chaturvedi, Pandit, and Srivastava 2006). In this work, the multiobjective function was converted into a single-objective function by introducing a novel price penalty factor function.

(Li et al. 2018) proposed a novel two stage approach to solve combined heat and power EED problem considering transmission losses and valve point effect. In the first stage of this process, Pareto optimal solutions were obtained through θ-dominance-based evolutionary algorithm (EA), while in the second stage, the obtained Pareto optimal solutions are separated into different clusters using fuzzy c-means and best compromise solutions were obtained using gray relation projection. The proposed technique yielded better results than NSGA-II and strength Pareto EA-2 (SPEA-2). An efficient fitness-based differential evolution (DE) (EFDE) algorithm and a constraint handling technique for dynamic EED (DEED) were proposed in (Shen et al. 2019). The EFDE approach offered better results when compared to other conventional DE variants. A hybrid NSGA-II with multiobjective particle swarm optimization (PSO) (MOPSO) was proposed in (Sundaram 2018) to solve the CEED problem of a power system. NSGA-II was applied in this work on best performing individuals while MOPSO is applied on lower ranked individuals and thus a balance between the exploration and the exploitation was maintained. The proposed technique performed much better than MOPSO and NSGA-II. A multiobjective multiverse optimization technique based on white hole, black hole, and worm hole concept of evolution of universe was utilized to solve the CEED and the combined heat and power dispatch of various power systems (Sundaram 2020). The proposed technique was applied on 10-, 40-, and 140- test systems. The results proved the superiority of this algorithm over other referred techniques. (Liang et al. 2018) proposed a multiobjective hybrid bat algorithm to formulate the CEED problem of IEEE 30-, 118-, and 300-bus test systems. The results show that this algorithm is very effective in solving the CEED problem even for large system.

The PSO technique was applied to solve the CEED problem considering various constraints like ramp rate constraint, prohibited operating zone (POZ) and nonsmooth fuel cost functions. The proposed technique was then compared with GA. The investigation proves superiority of PSO over GA (Gaing 2003).

An opposition based gravitational search algorithm (GSA) was proposed (Shaw, Mukherjee, and Ghoshal 2012) to solve the CEED problem while taking into consideration the effects of ramp rate

limit (RRL) and POZ. The technique reported in this work was found to be superior to the classical PSO technique.

Spiral optimization algorithm was applied to solve the CEED problem of a power system. Price penalty factor was considered to convert the multiobjective function into single-objective function. The technique was tested on 3-, 6-, and 40-generating unit systems (Benasla, Belmadani, and Rahli 2014). A new multi-objective cross entropy algorithm based on decomposition was proposed to solve a multiobjective optimization model for wind/hydro/thermal/photovoltaic power system in (Wang et al. 2020). The proposed algorithm was tested on IEEE 30- and 118-bus test power systems. Simulation results demonstrated that the algorithm reduced the fuel cost and the emissions effectively without violating any constraints.

A techno economic model is utilized to assess the levelized cost of electricity (LCOE) of an integrated biomass liquefaction plant with solar PV operating in Queensland, Australia. The study shows that the LCOE of the proposed system is comparable to that of solar PV and lithium-ion batteries (Perkins 2020). (Khan, Sidhu, and Gao 2016) proposed a new mixed integer binary optimization technique for a hybrid power system consisting of thermal and solar units. The proposed technique was verified on an IEEE 30-bus system.

The CEED problem is formulated using cubic criterion function and solved using Quantum mechanics based PSO technique. The proposed technique produces much better result compared to Lagrange relaxation, PSO, quantum bat algorithm and simulated annealing techniques (Mahdi et al. 2018). A similar problem of dispatch was solved using a quantum inspired tidal firefly technique. The proposed technique is a hybridization of tidal firefly algorithm with quantum mechanics principle. The proposed technique was able to optimize the problem efficiently and produced best results (Bodha, Yadav, and Mukherjee 2020). A differential evolution crossover quantum PSO is proposed (Xin-gang et al. 2020a). It uses a differential evolution operator to improve the global search ability of quantum PSO. The technique efficiently solves the CEED problem of an IEEE 30-bus system and delivers much better results than QPSO.

A DE based QPSO (Xin-gang et al. 2020b) was introduced to solve the CEED problem of a microgrid. The simulation results prove that the improved QPSO is more effective in solving the microgrid dispatching problem than QPSO and quantum-behaved PSO with collaborative attractors (CLQPSO). In (Xue et al. 2019), a time variable inertia weight (TVIW) strategy is utilized to update the velocity of the particle. PSO is integrated to provide GSA with memory. The proposed TVIW-GSA-PSO technique is applied to optimize the penalty parameter C and kernel function parameter of a support vector machine.

1.2. Research gap

The area of CEED of a hybrid solar thermal system is not much explored by researchers. Also, there is hardly any work that aims at renewable energy project development in India. This manuscript proposes a novel quantum inspired time varying inertia weight GSA PSO (QI-TVIW-GSA-PSO) algorithm to solve the CEED problem of a hybrid solar thermal system considering various constraints. The integration of any EA with quantum mechanics enhances its capability, as in quantum space the particles satisfy the basic notion of quantum mechanics that there is no restriction in the movement of objects in quantum space and the constraint of the search space is removed. Objects can move freely in the entire feasible solution space to find the optimal solution (Xin-gang et al. 2020b). The quantum inspired techniques explores the search space with a smaller population size and exploit global solution within a short span of time (Meng et al. 2010). An adaptive contraction expansion factor is also introduced in this manuscript that provides better exploration capability to the algorithm.

The simulation results are carried out on a six-generator system operating in New Delhi, India. The load data is taken from Tata Power Delhi Distribution Limited (Tata Power-DDL), India, a joint venture between Tata Power and the Government of National Capital Territory (NCT) of Delhi, India

for 3rd May 2020 (Centre, delhi state load dispatch, n.d.). The date has been arbitrarily chosen. The solar data are taken from solar simulator, designed and developed by Indian Institute of Technology Bombay (IIT Bombay) (Bombay, IIT, n.d.) for New Delhi region of India. In the present work, static combined economic emission dispatch (SCEED) is carried out for full and reduced solar radiation, while dynamic combined economic emission dispatch (DCEED) is considered over a period of 24 h. The technique is also applied to solve the CEED problem of 10 units, 2000 MW system and results are compared with six other established techniques.

1.3. Contribution of the present paper

The main contributions of the current work are presented below.

(a) A mixed integer binary programming problem is formulated that is highly constrained. A decomposition configuration is used in which the original problem is divided into two sub-problems.
(b) The proposed technique is an amalgamation of quantum mechanics and quantum computing principles with a variant of hybrid GSA-PSO. The quantum inspired TVIW-GSA-PSO (QI-TVIW-GSA-PSO) and quantum binary TVIW-GSA-PSO (QB-TVIW-GSA-PSO) techniques are exploited to find the joint optimized solution.
(c) An adaptive contraction expansion factor is introduced which ensures better exploration of the algorithm.
(d) The proposed framework consists of 13 solar PV panels of different characteristics unlike in the previous work where similar type of solar PV panels were considered.
(e) Constraint handling mechanism of a hybrid solar thermal system for CEED is demonstrated.

1.4. Layout of the present paper

The remainder of the article is summarized as follows: Section 2 deals with the mathematical formulation of the SCEED and the DCEED problem of a hybrid solar thermal plant. Section 3 discusses about GSA, TVIW-GSA-PSO optimization techniques and the novel binary TVIW-GSA-PSO. In Section 4, a novel QI-TVIW-GSA-PSO and QB-TVIW-GSA-PSO algorithms for mixed integer Optimization Problem (MIOP) are presented. Section 5 deals with results and discussions. Finally, a brief conclusion and future scope of work are highlighted in Section 6.

2. Problem formulation

SCEED and DCEED of a thermal power plant linked with solar PV cell are discussed. Following sections shows mathematical formulation aspect.

2.1. Mathematical formulation of hybrid SCEED

The CEED problem consisting of two objective functions, the fuel cost and the emission objective, is formulated as shown in (1)

$$MinC = \sum_{i=1}^{N} FC_i(P_i) + EC_i(P_i) \tag{1}$$

where C is the objective function to be minimized, $FC_i(\mathrm{P_i})$ represents the fuel cost (\$/h) and $EC_i(\mathrm{P_i})$ is the emissions in kg/h of the ith generating unit, respectively. The aim is to minimize the objective function subjected to various constraints as laid down in the succeeding subsections.

2.2. Equality constraint

The power balance equation is governed by (2)

$$\sum_{i=1}^{N} P_{gi} - (P_{loss} + P_{demand}) = 0 \tag{2}$$

where P_{gi} is the total power developed by the ith unit, P_{loss} is the transmission line loss, P_{demand} is the load demand, and N is the total number of generators.

2.3. Inequality constraint

The generator output should lie between its minimum and maximum generating values, as given by (3)

$$P_{igenmin} \leq P_{igen} \leq \mathrm{P_{igenmax}} \tag{3}$$

where $P_{igenmin}$ and $P_{igenmax}$ are minimum and maximum power limit of the ith generator, respectively. $FC_i(P_i)$, $EC_i(P_i)$ in Eq. (1) and P_{loss}in (2) can be formulated as follows (Chatterjee, Ghoshal, and Mukherjee 2012):

Fuel cost in \$/h for the ith generating unit ($FC_i(P_i)$) is given by (4)

$$FC_i(P_i) = A_i P_{igen}^2 + B_i P_{igen} + C_i + \left|D_i * \sin(E_i * \left(P_{igenmin} - P_{igen}\right)\right| \tag{4}$$

where A_i,B_i, and C_i represent the fuel cost coefficient of the ith generating unit. Opening of steam valves in thermal power plant causes sudden increase in observed losses, known as valve point loading. This can be modeled as ripples in the fuel cost function using *sine* terms. The fuel cost coefficients D_i and E_i are used to incorporate valve point effect of the ith generator. The valve point effect makes the fuel cost function highly nonlinear.

Emission/kg of the ith generating unit ($EC_i(P_i)$) is presented in (5)

$$EC_i(P_i) = alpha_i P_{igen}^2 + beta_i P_{igen} + gamma_i + L_i * \exp\left(Z_i * P_{igen}\right) \tag{5}$$

where $alpha_i$,$beta_i$,$gamma_i$,L_i, and Z_i are the emission coefficients of the ith generating unit. The transmission line losses can be calculated using Kron loss relation as shown in (6)

$$P_{loss} = \sum_{i=1}^{N}\sum_{j=1}^{N} P_{igen} B_{ij} P_{jgen} \tag{6}$$

where B is the loss coefficient matrix.

It may be noted that (1) can be converted to a single-objective function using price penalty factor h (Güvenç et al. 2012). After the inclusion of price penalty factor, the complete objective function can be defined as (7)

$$\begin{aligned} MinC = \sum_{i=1}^{N} \Big(& A_i P_{igen}^2 + B_i P_{igen} + C_i + \left|D_i * \sin(E_i * \left(P_{igenmin} - P_{igen}\right)\right| \Big) \\ & + h_i (alpha_i P_{igen}^2 + beta_i P_{igen} + gamma_i + L_i \times \exp\left(Z_i \times P_{igen}\right) \end{aligned} \tag{7}$$

where h_i is given as in (8).

$$h_i = \frac{A_i P_{igenmax}^2 + B_i P_{igenmax} + C_i + \left|D_i * \sin(E_i * \left(P_{igenmin} - P_{igenmax}\right)\right|}{alpha_i P_{igenmax}^2 + beta_i P_{igenmax} + gamma_i + L_i * \exp\left(Z_i * P_{igenmax}\right)} \tag{8}$$

The power generated by a solar plant is represented by (9) (Shilaja and Ravi 2017)

$$P_{solar} = P_{rated}\left\{1 + \left(t_{Reference} - t_{Ambient}\right) * \tau\right\} * \frac{SR_i}{1000} \tag{9}$$

where P_{rated} is the rated power, $t_{Reference}$ is the reference temperature, $t_{Ambient}$ is the ambient temperature, τ is the temperature coefficient, and SR_i is the incident solar radiation of the ith solar plant.

The solar share and the solar cost of a solar plant are, respectively, given by (10) and (11)

$$Solar\ share = \sum_{j=1}^{M} P_{solarj} \times SW_j \tag{10}$$

$$Solar\ cost = \sum_{j=1}^{M} puCost_j \times P_{solarj} \times SW_j \tag{11}$$

where M is the total number of solar plant, $puCost_j$ is the per unit cost of the jth solar plant, SW_j is the switching state of the jth solar plant. If it is taken as 0, it means that the solar plant is kept OFF, while if the value is 1, it is turned ON.

Another objective is to maximize the benefits of installed solar power by raising number of participating solar units and to reduce the likelihood of solar failure by increasing the solar share. (Khan et al. 2015). Therefore, with solar generation included in dispatch, the updated objective function can be represented as in (12)

$$\begin{aligned} MinC = & \sum_{i=1}^{N}\left(A_i P_{igen}^2 + B_i P_{igen} + C_i + \left|D_i * \sin(E_i * \left(P_{igenmin} - P_{igen}\right)\right|\right) \\ & + h_i(alpha_i P_{igen}^2 + beta_i P_{igen} + gamma_i + L_i \times \exp\left(Z_i \times P_{igen}\right) + \sum_{j=1}^{M} puCost_j \times P_{solarj} \times SW_j \\ & + \mathrm{K}\left(\sum_{j=1}^{M} P_{solarj} - \sum_{j=1}^{M} P_{solarj} \times SW_j\right) \end{aligned} \tag{12}$$

subject to

$$P_{demand} + P_{loss} - \sum_{i=1}^{N} P_{gi} - \sum_{j=1}^{M} P_{solarj} \times SW_j = 0 \tag{13}$$

$P_{min} \leq P_{gen} \leq P_{max}$(14)

where, P_{min} and $\mathrm{P_{max}}$ are the minimum and maximum generating limit of the generators respectively.

$$\sum_{j=1}^{M} P_{solarj} \times SW_j \leq 0.3 \times P_{demand} \tag{15}$$

where K is a constant that is employed to make the last term of (12) reconcilable with other terms of the equation under reference.

2.4. Mathematical formulation of dynamic CEED with solar power

DCEED refers to scheduling of generators for minimum cost of operation for the entire dispatch period. The Ramp Rate Limit (RRL) is considered in this problem. To incorporate the DCEED problem, the mathematical formulation of (12) may be shown as represented in (16).

$$MinC_t = \sum_{t=1}^{T}\sum_{i=1}^{N}\left(A_i\left(P_{igen}^t\right)^2 + B_iP_{igen}^t + C_i + \left|D_i * \sin(E_i * \left(P_{igenmin} - P_{igen}^t\right)\right|\right) + h_i(alpha_i\left(P_{igen}^t\right)^2 + beta_iP_{igen}^t + gamma_i + L_i \times \exp\left(Z_i \times P_{igen}^t\right) + \sum_{j=1}^{M} puCost_j \times P_{solarj} \times SW_j + \mathrm{K}\left(\sum_{j=1}^{M} P_{solarj}^t - \sum_{j=1}^{M} P_{solarj}^t \times SW_j^t\right)\right) \tag{16}$$

The RRLs represent the extent to which a generation can be increased or decreased (Koltsaklis, Dagoumas, and Mladenov 2020). The power generation of the generating unit is limited by the RRLs. Ramp-up and ramp-down constraints may be represented, in order, as in (17) and (18)

$$P_{igen}^t - P_{igen}^{t-1} \le RU_i \tag{17}$$

$$P_{igen}^{t-1} - P_{igen}^t \le RD_i \tag{18}$$

where RU_i and RD_i are the ramp-up and the ramp-down rate of the ith generating unit, respectively.

The minimum and maximum generating limit when ramp rates are considered is shown in (19).

$$Max(P_{igenmin,}RU_i - P_{igen}^t) \le P_{igen}^t \le Min(P_{igenmax,}P_{igen}^{t-1} - RD_i) \tag{19}$$

The power balance constraint for DCEED can be written as in (20).

$$P_{demand}^t + P_{loss}^t - \sum_{j=1}^{M} P_{solarj}^t \times SW_j^t = 0 \tag{20}$$

The total solar share should not be more than 30% of the total load. This can be shown as in (21).

$$\sum_{j=1}^{M} P_{solarj}^t \times SW_J^t \le 0.3 \times P_{demand}^t \quad SW_j^t \in \{0,1\} \tag{21}$$

2.5. Constraint handling

Constraint handling mechanism ensures that no power unit exceeds its defined limits. For effective implementation of this problem, Generator number 1 is chosen as slack generator to satisfy equality constraints. The power output of this unit can be computed by updating the energy balance equation represented in (2), while the updated equation is shown by (22).

$$P_s + \sum_{\substack{i=1 \\ i \ne s}}^{N} P_{gi} + \sum_{j=1}^{M} P_{solarj} \times SW_j = (P_{loss} + P_{demand}) \tag{22}$$

Considering transmission losses, Eq. (22) can be written as in (23).

$$P_s + \sum_{\substack{i=1 \\ i \ne s}}^{N} P_{gi} + \sum_{j=1}^{M} P_{solarj} \times SW_j =$$

$$\sum_{\substack{i=1\\ i\neq s}}^{N}\sum_{\substack{J=1\\ J\neq 1}}^{N} P_i B_{ij} P_j + \sum_{\substack{j=1\\ j\neq s}}^{N} P_j(B_{js+}B_{sj})P_s + B_{ss}P_s^2 + \sum_{\substack{i=1\\ i\neq S}}^{N} B_{i0}P_i + B_{S0}P_s + B_{00} + P_{demand} \tag{23}$$

Eq. (23) can be written as in (24)

$$X_1 P_s^2 + Y_1 P_s + Z_1 = 0 \tag{24}$$

where

$$X_1 = B_{ss} \tag{25}$$

$$Y_1 = \sum_{\substack{j=1\\ j\neq s}}^{N} P_j(B_{js+}B_{sj}) + B_{S0} - 1 \tag{26}$$

$$Z_1 = P_{demand} + \sum_{\substack{i=1\\ i\neq s}}^{N}\sum_{\substack{j=1\\ j\neq s}}^{N} P_i B_{ij} P_j + \sum_{\substack{i=1\\ i\neq S}}^{N} B_{i0}P_i - \sum_{\substack{i=1\\ i\neq s}}^{N} P_{gi} - \sum_{j=1}^{M} P_{solarj} \times SW_j \tag{27}$$

The positive roots of this equation can be calculated by (28).

$P_s = \frac{-Y_1 \pm \sqrt{Y_1^2 - 4X_1Z_1}}{2X_1}$ where $(Y_1^2 - 4X_1Z_1) \geq 0$ (28)

Only positive roots of (28) are considered. If roots are out of bound, the slack generator power $P_s$is set to its minimum limit, P_s^{min}. If both values are within limits, then the root with the minimum value is considered. Calculate the slack generation gap (P_{slack_gap}) using (29).

$P_{slack_gap} = \sum_{i=1}^{N} P_{gi} - (P_{loss} + P_{demand})$ (29)

Repeat this process until the slack generation gap given by (29) reduces to zero.

3. Optimization techniques

The CEED of a hybrid thermal solar plant is MIOP in nature. Decision variables for the thermal generating units are continuous whereas decision variables for solar units are binary. In order to solve this problem, QI-TVIW-GSA-PSO and QB-TVIW-GSA techniques are used.

3.1. Review of GSA

The GSA is a population based optimization algorithm that is based on Newton's gravitational principles. In GSA, agents are considered to be objects and performance is measured by their mass. All these bodies are governed by the gravitational law and all other bodies move toward the body with heavier mass. The heavier mass corresponds to good solution and travels slower than bodies having lighter mass. This property ensures better exploitation of the algorithm (Rashedi, Nezamabadi-pour, and Saryazdi 2009).

Every agent is characterized by four parameters *viz.* active gravitational mass, passive gravitational mass, inertia mass and position. The mass represents a potential solution and the algorithm converges by proper tuning of gravitational and inertial mass. The mass is attracted by the heaviest mass and an optimal solution is obtained with the passage of time (Yazdani, Nezamabadi-Pour, and Kamyab 2014).

GSA can be viewed as an isolated system of masses. It is like a small universe of particles that follow Newton's force and motion rules. These masses obey the following two principles:

(a) Law of gravity – Particles attract each other and the gravitational force between them is directly proportional to the product of their masses and inversely proportional to square of distance (R) between them. In GSA, instead of square of the distance between particles, only R is taken considering it provides better solution.
(b) Law of motion – The current velocity of the mass is equal to the sum of its previous velocity and acceleration. The acceleration is given by force acting on the mass divided by its inertial mass.

Consider a system with n objects where the location of the ith object is defined as in (30)

$A_i=(a_i^1, \ldots ., a_i^d, \ldots . a_i^n)$ for $i = 1,2, \ldots ., n$ (30)

where a_i^d represents the position of the ith object at the dth dimension.

The force acting on the ith mass from the jth mass at a specific time is given by the following equation

$F_{ij}^d(t) = \mathrm{G}(t) \times \frac{M_{passivei}(t) \times M_{activej}(t)}{d_{ij}(t)+e} \times (a_j^d(t) - a_i^d(t))$ (31)

where $M_{passivei}(t)$ is the passive gravitational mass of the object i at time (t) while $M_{activej}(t)$ is the active gravitational mass of the jth object at time t, G(t) denotes the gravitational constant at time t, e is a constant and $d_{ij}(t)$ is the Euclidian distance between two objects i and j. It can be calculated using (32).

$$d_{ij}(t) = a_i(t), a_j(t)_2 \tag{32}$$

To give a stochastic characteristic to the GSA, the total force acting upon the ith object in the dth dimension is the weighted sum of the dth components of force exerted from other objects and is given by (33)

$$F_i^d(t) = \sum_{j=1, j\neq i}^{N} rand \times F_{ij}^d \tag{t)(33}$$

where *rand* is a random number that lies in [0, 1]. According to the law of motion, the acceleration of the ith object at time t in the dth dimension is given by (34)

$Acl_i^d(t) = \frac{F_i^d(t)}{M_{inertiali}(t)}$ (34)

where $M_{inertiali}$(t) is the inertial mass of the ith object. The velocity and the position of the ith object at time t is given by (35) and (36), respectively (Duman et al. 2012)

$$Vel_i^d(t+1) = rand_i \times Vel_i^d(t) + Acl_i^d(t) \tag{35}$$

$$a_i^d(t+1) = a_i^d(t) + Vel_i^d(t+1) \tag{36}$$

where $rand_i$ is the random number between [0, 1]. The random number provides a randomized property to the search. Gravitational constant is initialized at the beginning and is reduced with time to provide better search capability. It is expressed as in (37)

$$\mathbb{G}(t) = \mathbb{G}(t_0) \times \left(\frac{t_0}{t}\right)^{\alpha}, \alpha < 1 \tag{37}$$

where $\mathbb{G}(t)$ is the gravitational constant at time t and $\mathbb{G}(t_0)$ is the gravitational constant at the first cosmic quantum interval of time t_0. The gravitational constant is taken into consideration while following (38)

$$\mathbb{G}(t) = \mathbb{G}(t_0) \times e^{-\tau\left(\frac{it}{itmax}\right)} \tag{38}$$

where $\mathbb{G}(t_0)$ is set to 100, τ is taken as 20, *it* and *itmax* are current and maximum number of iteration, respectively (Shaw, Mukherjee, and Ghoshal 2012).

The gravitational constant is the modulus that controls the impact of gravitational force. The balance between the exploration and the exploitation depends upon the value of gravitational constant. A large value of gravitational constant enables global exploration while a smaller value of gravitational constant makes the algorithm move toward fine tuning of the current search area. This decreasing value is achieved by using the exponential function described in (38).

The gravitational and the inertial masses are calculated by evaluating fitness function. A heavier mass represents an efficient solution. Active mass, passive mass and inertial mass are though conceptually different but no experiment has ever unambiguously shown any difference between them. The theory of general relativity is based on the assumption that inertial mass and passive gravitational mass are equivalent. This is also known as weak equivalence principle. The strong principle of equivalence, a key stone to Einstein theory of general relativity, is also laid on the principle that active gravitational mass and inertial mass are equivalent (Rashedi, Nezamabadi-pour, and Saryazdi 2009). Hence, the masses are kept equal as expressed in (39)

$M_{activei} = M_{passivei} = M_{inertiali}$for i = 1,2,3 . . . .n. (39)

$$m_i(t) = \frac{fitness_i(t) - worst(t)}{finest(t) - worst(t)} \tag{40}$$

$$M_i(t) = \frac{m_i(t)}{\sum_{i=1}^{N} m_i(t)} \tag{41}$$

where $fitness_i(t)$ is the fitness value of the ith object at time t.

$$Worst(t) = \max_{j\in\{1,\ldots,n\}} fitness_j(t) \tag{42}$$

$$\text{fitness}(t) = \min_{j\in\{1,\ldots,n\}} fitness_j(t) \tag{43}$$

To create a balance between exploration and exploitation, only objects with heavier mass are considered to attract other masses.

An optimization algorithm must use exploration at the beginning and with passage of time, explorations should be reduced while the exploitation should step in. To improve the performance of GSA, a term w_{best} is introduced to modify (33). The new modified equation becomes as in (44)

$$F_i^d(t) = \sum_{j\in wbest, j\neq i} rand_j F_{ij}^d(\mathrm{t}) \tag{44}$$

where w_{best} is the set of first w objects with heavier weight or best fitness values.

3.2. TVIW-GSA-PSO algorithm

There are some drawbacks in GSA like (a) it suffers from slow convergence, (b) position update depends only upon the current position and (c) it does not take into account any exchange of information between particles. These issues can be addressed by integrating PSO technique with GSA and the new velocity update equation of the integrated GSA-PSO can be written as in (45) (Mirjalili and Hashim 2010).

$$Vel_i^d(t+1)' = \sigma \times Vel_i^d(t)' + K_1 \times rand \times \left(Acl_i^d(t)\right) + K_2 \times rand \times \left(Globest - a_i^d(t)\right) \tag{45}$$

Here, K_1 and K_2 are constants with values lying in [0, 1]. Here, *Globest* is the global best solution obtained. The symbol σ is the global inertia weight, it is used to fine tune the algorithm between the global search and the local search. The larger value of σ causes the algorithm to speed up but the inaccuracy increases while the exploration is decreased. On the other hand, a very small value of σ

unnecessarily slows down the convergence. Eq. (45) is divided into three parts, the first part is the velocity of objects, the second part is object acceleration and the third part shows the exchange of data between particles. The third part defines the trajectory of the object.

To create a balance between the exploration and the exploitation, the inertia weight σ is calculated by employing TVIW strategy stated in (46)

$$\sigma(t) = \sigma_{max}\left(\frac{\sigma_{min}}{\sigma_{max}}\right)^{\frac{it}{itmax}} \tag{46}$$

where σ_{max} is the maximum inertia weight, while σ_{min} is the minimum inertia weight. Here, *it* and *itmax* are current iteration and maximum iteration, respectively. At $t = 0$, σ(t) =σ_{max} and at t = itmax, σ (t) =σ_{min}. This ensures a very strong global search capability at the starting and, consequently, the algorithm converges quiet rapidly and during the final count of the iteration, the value of inertia weight is reduced to enhance the local search capability. This algorithm is referred as TVIW-GSA-PSO (Xue et al. 2019). The position update equation is governed by (47).

$$a_i^d(t+1)' = a_i^d(t)' + Vel_i^d(t+1)' \tag{47}$$

3.3. Binary TVIW-GSA-PSO algorithm

The position update rule shown in (47) updates the position for power in thermal generating unit. For solar generating unit, the decision variable is 0 (corresponds to turned OFF) or 1 (means turned ON). So, for updating the state of solar unit, binary TVIW-GSA-PSO is utilized (Ji et al. 2014). In binary optimization, the position update means switching to a "0" or "1" state. The switching should be done according to the mass velocity. A larger change in velocity means that the current position of the particle is not good and it requires a major change from its current position, while if the velocity is very small, it indicates that the particle is very nearer to its optimal position and requires only a minor change from its position. The function $|\tanh Vel_i^d(t+1)'|$ is bounded between [0, 1] and satisfies above mentioned requirements (Rashedi, Nezamabadi-Pour, and Saryazdi 2010). The position update or change in the state of the solar power unit is governed by the following equation.

$$f(x) = a_i^d(t)' if rand \tanh Vel_i^d(t+1)' a_i^d(t)', and\ otherwise \tag{48}$$

The next state is dependent on the velocity of the particle given by (45). The steps involved in TVIW-GSA-PSO for mixed integer-based CEED problem is summarized in Algorithm 1.

4. Quantum inspired optimization techniques

The TVIW-GSA-PSO addresses all shortcomings of the GSA and shows excellent performance. The algorithm mimics the PSO algorithm and suffers from some drawbacks. In TVIW-PSO-GSA, there is certain attraction potential energy due to which particles keep on aggregating and move in a particular trajectory. When the TVIW-GSA-PSO is applied though the principle of quantum mechanics, constraint of the search space is removed, the objects can move freely in the entire feasible solution space to find the optimal solution.

It is assumed that, in a quantum space, the particles satisfy the basic notion of quantum mechanics. There is no restriction in the movement of the objects in quantum space. Lack of a defined trajectory in quantum search space means the velocity and the position of objects cannot be calculated simultaneously. Therefore, the wave function ψ is used to describe the state of the particle.

In quantum mechanics, the time-dependent Schrodinger equation (TDSE) deals with the time and the position (Mahdi et al. 2019) (see (49) and (50))

$$i\frac{\partial}{\partial t}\Psi(x,t) = \hat{H}(x)\Psi(x,t) \tag{49}$$

$$\hat{H}(x) = -\frac{^2}{2m_a} \quad ^2 + P(x) \tag{50}$$

$$^2 = \frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2} + \frac{\partial}{\partial z^2} \tag{51}$$

where $\Psi(f,t)$ is the wave function of an object in quantum space expressed in terms of position (x) and time (t), $\hat{H}$ is Hamiltonian operator, m_a is the mass of the particles, $P(x)$ is the potential energy and ^{2} is the Laplacian.

The wave function $(\Psi(x,t))$ of a particle in quantum system does not have any physical significance, but the square of its amplitude gives the probability of finding a particle. It can be expressed by (52).

$$\int_{-\infty}^{+\infty} |\Psi(x,t)|^2 dxdydz = 1 \tag{52}$$

In the proposed technique, square potential well is chosen. Assume that the quantum space is having dimension d and n is the number of particles. Location of the ith particle is $X_i = \left(x_{i1}, x_{i2}, .., x_{id}\right)$, w_{id}^t is the local attractor of particle i in dimension d at iteration t. The local attractor is the best position at iteration t and is determined by using the particle position update stated in (47). The particle position update equation using Monte Carlo stochastic simulation can be written as in (53)

$$x_{id}^{t+\Delta t} = w_{id}^t \pm \frac{C}{2}\ln(1/r \tag{53}$$

where r is a random number in [0, 1] and c depends on particles' current and the best position and can be calculated using (54)

$$c = 2.\delta.\left|W_{id}^t - x_{id}^t\right| \tag{54}$$

where

$$W_{id}^t = \frac{1}{N}\sum_{i=1}^{N} w_{id}^t \tag{55}$$

The updated equation is now obtained by (56)

$$x_{id}^{t+\Delta t} = w_{id}^t \pm \delta.\left|W_{id}^t - x_{id}^t\right|.\ln(1/r) \tag{56}$$

where $\delta$is the adaptive contraction expansion factor. It is tuned such that it has a lower initial value and increases with the number of iterations. The adaptive contraction expansion factor is formulated as in (57)

$\delta = (1 \rightarrow 0.5)\frac{it_{max'} - it'}{it_{max'}} + \sin(\theta)$ (57)

where θ is the incremental angle that increases from 0 to 0.5235 rad, $it_{max'}$and $it'$are the maximum iteration and the current iteration, respectively.

4.1. Quantum inspired binary TVIW-GSA-PSO

The position update presented in (56) is for optimization of thermal generating units and generates continuous values. For the solar unit, the decision variables are binary. So, these variables are updated using QB-TVIW-GSA-PSO. A rotation transformation matrix (Ji et al. 2014) is used to update the particle position. It is given by (58)

$$R(\Delta\phi) == \begin{bmatrix} cos(\Delta\phi) & -sin(\Delta\phi) \\ \sin(\Delta\phi) & cos(\Delta\phi) \end{bmatrix} \tag{58}$$

where R() is the rotation transformation function and Δϕ is the rotation factor.

The rotation factor Δϕ of a particle i in dimension d at iteration t is given by (59)

$$\Delta\phi_{id}^{t} = \sum_{j \neq i} (\phi \times \eta_{id}^{,t} \times \left(a_{id}^{t} - a_{jd}^{t}\right)) \tag{59}$$

where the rotation angle ϕ decreases from $\phi_{maximum}$ to $\phi_{minimum}$ and is given by (60)

$$\phi = \phi_{maximum} - (\phi_{maximum} - \phi_{minimum}) \frac{it}{it_{max}} \tag{60}$$

where it and it_{max} are current iteration number and maximum iteration number, respectively.

The variable $\eta_{id}^{,t}$ of (59) depends on λ_{id}^{t} which acquires the value 0 or 1 depending on the condition as shown in (61).

$$\lambda_{id}^{t} = \{ \begin{matrix} 1, if M_{activej}(t) > M_{passivej}(t) and d_{ij}^{h}(t) \leq \\ 0, otherwise \end{matrix} \tag{61}$$

The variable $\eta_{id}^{,t}$ is opposite of the value of λ_{id}^{t} if $\min(C) > Globest$, else its value is same as expressed in (62)

$$\eta_{id}^{,t} = \left\{ \overline{\lambda_{id}^{t}}, if \min(C) Globest \lambda_{id}^{t}, otherwise \right. \tag{62}$$

where d_{ij}^{h} is the hamming distance between the active mass and the passive mass, δ is number of dimension considered. Here, min (C) is the objective function to be minimized. The position is updated through (63).

$$\begin{bmatrix} alpha_{i,d}^{t+1} \\ beta_{i,d}^{t+1} \end{bmatrix} = \mathrm{R}(\Delta\phi_{id}^{t}) \times \begin{bmatrix} alpha_{i,d}^{t} \\ beta_{i,d}^{t} \end{bmatrix} \tag{63}$$

The position update for binary variables is governed by (64)

$$u_{i,d}^{t+1} = \begin{cases} 1, if rand < \left| beta_{i,d}^{t+1} \right|^2 \\ 0, otherwise \end{cases} \tag{64}$$

where $\left| alpha_{i,d}^{t+1} \right|^2$ and $\left| beta_{i,d}^{t+1} \right|^2$ give the probability that the position $u_{i,d}^{t}$ will be in "0" state and "1" state, respectively.

The optimized value of thermal units is calculated using (56) while the operating state of solar unit is calculated using (64). The step by step algorithm for QI-TVIW-GSA-PSO for MIOP is summarized in Algorithm 2.

5. Results and discussions

The proposed technique has been implemented to solve SCEED and DCEED problem. The model is simulated on MATLAB-2016a using an Intel Celeron 1.10 GHz dual core CPU with 4GB of RAM. The proposed algorithm is run for 120 iterations. For each case, 20 runs are considered and minimum values of obtained results are considered.

5.1. Test case-I

This case considers a 10-unit generating system considering valve point effect and transmission line losses. The fuel cost coefficient, emission coefficient, and transmission line loss coefficient are taken from (Ali and Abd Elazim 2016). A load of 2000 MW is considered and results are compared with multi-objective DE (MODE), NSGA-II, Pareto DE (PDE), SPEA-II (Basu 2011), flower pollination algorithm (FPA) (Abdelaziz, Ali, and Abd Elazim 2016), and mine blasting algorithm (MBA) (Ali and Abd Elazim 2016). Results are tabulated in Table 1. From Table 1, it may be inferred that QI-TVIW-GSA-PSO is a capable algorithm and has optimized the problem successfully without violating any constraint. The proposed algorithm gives the lowest fuel cost out of all compared techniques. It gives a saving of 1254$ over MODE, 1309$ over NSGA-II and 1140$ over FPA.

5.2. Test case-II

In this case, SCEED is conducted on a 6 generator system with 13 solar PV cell operating in New Delhi region of India for full and reduced solar radiation. A 6-unit IEEE 30-bus system is considered for the standardization of the problem. The fuel cost coefficient data and emission coefficient data are given, respectively, in Tables A1 and Table A2 of Appendix Section (Manteaw and Odero 2012). Figure 2 shows the solar power share (in MW) for full and reduced solar radiation.

Solar power rating and per unit solar cost are shown in Table A3 of Appendix Section (REN21 2018). The hourly load data is taken from Tata Power-DDL, a joint venture between TATA power and Government of the NCT of Delhi and is responsible for handling the distribution of power in north and north-west parts of Delhi (Centre, delhi state load dispatch, n.d.) for 3rd May 2020 and the solar radiation data are taken from IIT Bombay, India solar thermal plant simulator (Bombay, IIT, n.d.) (refer Table A4 of Appendix Section). This toolkit is used for simulating the direct solar radiation of New Delhi region on 3rd of May 2020.

Table 2 shows SCEED results for full and reduced solar radiation. It can be seen that, for the particular day, panels have received a low solar radiation owing to cloudy weather. The pattern [1 1 1 1 1 1 1 1 1 1 1 1 1] represents a full solar radiation and all cells are turned ON, while the pattern [0 1 0 1 1 1 1 0 1 1 0 1 1] represents reduced solar radiation. Solar shares referring to full and reduced radiation are 200.1546 MW and 143.5502 MW, respectively. With reduced solar radiation, the solar cost also reduces but fuel cost and emission cost increases. A 30% reduction in solar share raises the total cost by 11%.

Table 3 shows the SCEED problem at time 11:00 h for a load of 705.73 MW. The solar power share for full radiation is 197.0452 MW and reduces to 138.8273 MW for cloudy weather. It may be noted that 30% reduction in solar radiation causes 10% increase in fuel cost, increases line losses and 12% increase in total cost. Figure 1 shows the solar power share for full and reduced solar radiation.

Table 1. Comparison of CEED results for a 10-unit system at 2000 MW load.

Power output (MW)	MODE	NSGA-II	PDE	SPEA-II	FPA	MBA	QI-TVIW-GSA-PSO
P_1	54.94	51.95	54.98	52.97	53.18	45.50	54.42
P_2	74.58	67.25	79.38	72.81	79.97	79.99	65.95
P_3	79.42	73.68	83.98	78.11	78.10	105.63	107.40
P_4	80.68	91.35	86.59	83.60	97.11	88.83	89.12
P_5	136.85	134.05	144.43	137.24	152.74	135.27	109
P_6	172.63	174.95	165.77	172.91	163.08	163.50	150.77
P_7	283.82	289.43	283.21	287.20	258.61	261.26	259.14
P_8	316.34	314.05	312.77	326.40	302.22	281.23	319.80
P_9	448.59	455.69	440.11	448.88	433.21	456.47	459.83
P_{10}	436.42	431.80	432.67	423.90	466.07	466.91	470
Fuel cost ($)	113484	113539	113510	113520	113370	112962	112230
Emission (lb)	4124.9	4130.2	4111.4	4109.1	3997.7	4077.8	4153.9
Losses (MW)	84.33	84.25	83.9	84.1	84.3	84.7	85.43
CPU (sec)	3.82	6.02	4.23	7.53	2.23	2.17	7.83

Table 2. SCEED result with integrated solar power for 699.95 MW load at 10:00 h.

Units	Power generation (MW)	
P_1	33.2761	41.4385
P_2	26.948	36.7375
P_3	89.8169	98.3217
P_4	90.3774	98.5944
P_5	135.3807	147.4950
P_6	132.5466	144.7497
Solar generation	**[1 1 1 1 1 1 1 1 1 1 1 1]**	**[0 1 0 1 1 1 1 0 1 1 0 1 1]**
Solar power share	200.5146	143.5502
Fuel cost ($/h)	27589	30313
Emission cost ($/h)	12616.8297	14655.24771
Solar cost ($/h)	40.0573	28.8468
Total cost ($/h)	40245.89	44997.09
Emission (kg/h)	262.5771	305.0063
Line losses (MW)	8.9103	10.9370

Table 3. SCEED result with integrated solar power for 705.73 MW load at 11:00 h.

Units	Power generation (MW)	
P_1	34.6000	42.9456
P_2	28.5336	38.5487
P_3	91.1947	99.8950
P_4	91.7092	100.1134
P_5	137.3453	149.7325
P_6	134.5261	147.0028
Solar generation	**[1 1 1 1 1 1 1 1 1 1 1 1 1]**	**[0 1 1 0 1 1 1 1 1 1 1 0 0]**
Solar power share	197.0452	138.8273
Fuel cost ($/h)	28026	30823
Emission cost ($/h)	12927.7660	15063.83357
Solar cost ($/h)	39.3643	27.7207
Total cost ($/h)	40993.13	45914.55
Emission (kg/h)	269.0482	313.5033
Line losses (MW)	9.2241	11.3354

At time 12:00 h, the solar radiation further decreases causing a drop in solar power share. At full solar radiation, the solar power share is 185.8428 MW while the solar power share during reduced radiation is 126.7110 MW. The reduction causes an increase in emission cost by 17% while the total cost increases by 12% as shown in Table 4. Line losses are also increased.

Table 5 shows the SCEED for 13:00 h. The solar power share for the full radiation case is 26.5% of the total load. The solar share becomes 19% when the solar PV cell experiences reduced solar radiation. The solar cost does not have a major impact on total operating cost.

Table 6 shows the SCEED for a load of 719.57 MW at 14:00 h. The solar power share is 27% of the total load. It reduces to 19% when the solar radiation is reduced. There is a total saving of 4887$/h when all the solar panels are operational.

At 15:00 h (see Table 7), the load as well as the solar radiation are increased. The solar power share is approximately 27% of the total load and, with reduced solar radiation; the solar power share reduces to 19%. The fuel cost increases by 10%. And the total cost increases by 11%. Line losses also increased by 30%.

The proposed QI-TVIW-GSA-PSO performs well in extremely constrained environment. The scheduling is within generating limit, the solar share is less than 30% of the total demand and losses are also considered. Figures 3 and 4 show that the convergence characteristic for the SCEED at full and reduced solar radiation. It may be observed that the algorithm converges rapidly for all the cases.

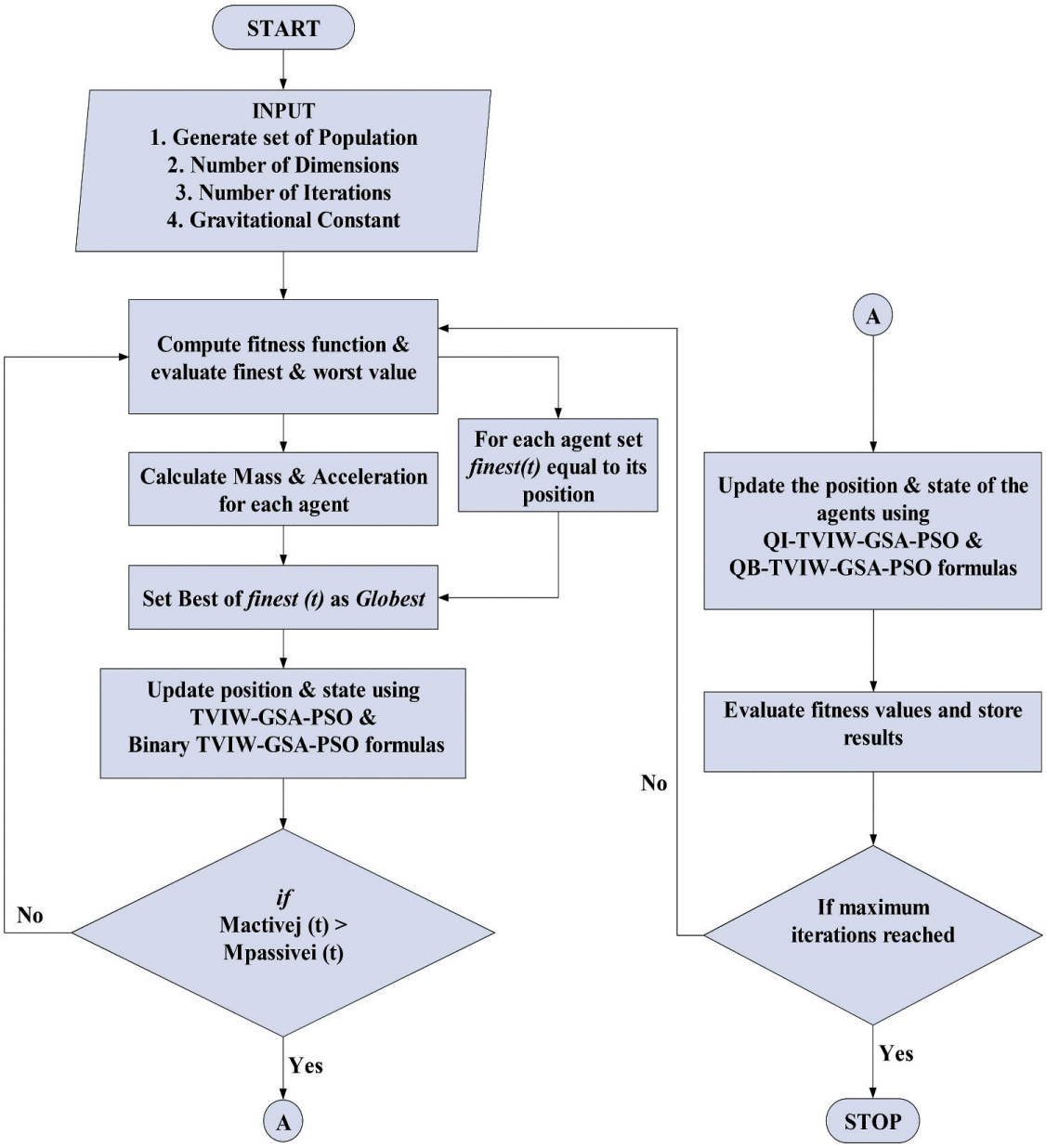


Figure 1. Flowchart of the proposed QI-TVIW-GSA-PSO technique.

5.3. Test case-III

In this case, ramp rate constraints are considered and the data for the same is given in Table A5 of Appendix Section. The data for solar units are same as that in Test Case-II, while the thermal data and the hourly load data are taken from (Elaiw, Xia, and Shehata 2012) and is presented in Table A6 of Appendix Section. The 13 solar plants are considered operating in New Delhi region of India with reduced solar radiation.

Table 8 shows the DCEED solution for 6-unit system with load demand shown in Table A6 of Appendix Section. Solar plants are considered to be operational from 10:00 h to 15:00 h for reduced solar radiation in New Delhi region of India. Hence, the solar radiation data remains same as that in SCEED problem. Ramp rate is considered for this case while line losses are neglected. The larger solar share results in reduced fuel cost, less emission and lesser emission cost. The cost/MW is around 20 $ and with solar integration, it further reduces to 16 . USD Also, it should be noted that the solar radiation received on the chosen day (*i.e.* 3rd May 2020) was much less than the usual radiation that

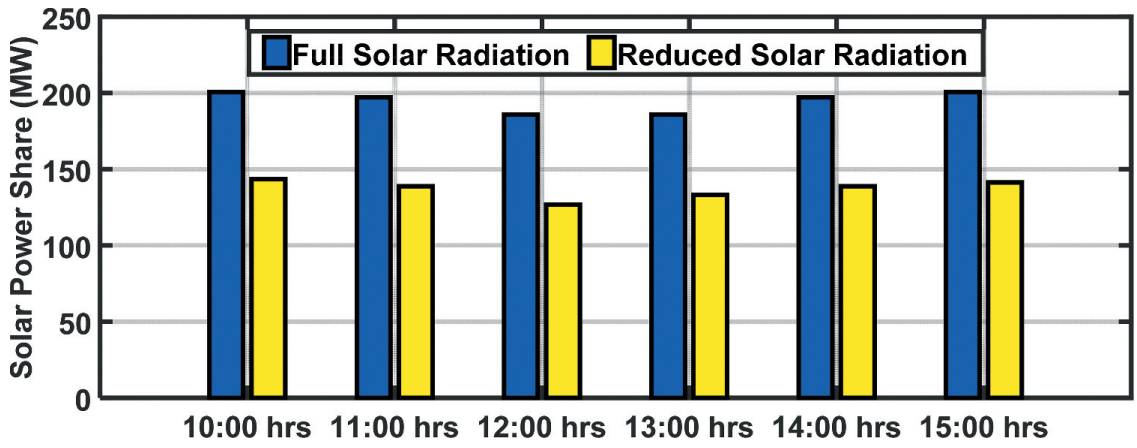


Figure 2. Solar power share for full and reduced solar radiation.

Table 4. SCEED result with integrated solar power for 697.58 MW load at 12:00 h.

Units	Power generation (MW)	
P1	35.0371	43.5150
P2	29.0573	39.2332
P3	91.6496	100.4895
P4	92.1488	100.6874
P5	137.9939	150.5778
P6	135.1795	147.8540
Solar generation	**[1 1 1 1 1 1 1 1 1 1 1 1 1 1]**	**[1 1 0 1 1 1 0 0 1 1 1 0 1]**
Solar power share	185.8428	126.7110
Fuel cost ($/h)	28171	31016
Emission cost ($/h)	12286.2298	14349.5451
Solar cost ($/h)	37.1263	25.1944
Total cost ($/h)	40494.36	45390.74
Emission (kg/h)	271.2192	316.7670
Line losses (MW)	9.3290	11.4878

Table 5. **S**CEED result with integrated solar power for 698.92 MW load at 13:00 h.

Units	Power generation (MW)	
P1	35.2290	42.7979
P2	29.2872	38.3711
P3	91.8494	99.7407
P4	92.3419	99.9645
P5	138.2786	149.5132
P6	135.4664	146.7820
Solar generation	**[1 1 1 1 1 1 1 1 1 1 1 1 1 1]**	**[0 1 0 1 1 1 1 0 1 1 0 1 1]**
Solar power share	185.8428	133.0466
Fuel cost ($/h)	28234	30773
Emission cost ($/h)	12329.6589	14163.556
Solar cost ($/h)	37.1263	26.7360
Total cost ($/h)	40600.7852	44963.292
Emission (kg/h)	272.1779	312.6613
Line losses (MW)	9.3752	11.2960

Table 6. SCEED result with integrated solar power for 719.57 MW load at 14:00 h.

Units	Power generation (MW)	
P1	36.5821	44.9327
P2	30.9091	40.9384
P3	93.2585	101.9705
P4	93.7036	102.1170
P5	140.2867	152.6828
P6	137.4894	149.9733
solar generation	**[1 1 1 1 1 1 1 1 1 1 1 1 1]**	**[1 1 0 0 1 1 0 1 1 1 1 0 1]**
Solar power share	197.0452	138.8273
Fuel cost ($/h)	28684	31499
Emission cost ($/h)	12640.2040	14723.5101
Solar cost ($/h)	39.3643	28.2357
Total cost ($/h)	41363.57	46250.75
Emission (kg/h)	279.0332	325.0223
Line losses (MW)	9.7047	11.8720

Table 7. SCEED result with integrated solar power for 753.37 MW load at 15:00 hr.

Units	Power generation (MW)	
P1	40.9300	49.4390
P2	36.1267	46.3652
P3	97.7912	106.6831
P4	98.0821	106.6642
P5	146.7402	159.3749
P6	143.9896	156.7096
Solar generation	**[1 1 1 1 1 1 1 1 1 1 1 1 1]**	**[1 1 0 1 0 1 0 1 1 1 1 0 1]**
Solar power share	200.5146	141.2716
Fuel cost ($/h)	30141	33046
Emission cost ($/h)	13689.0439	15967.33494
Solar cost ($/h)	40.0573	28.7328
Total cost ($/h)	43870.10	49042.07
Emission (kg/h)	302.1864	352.4798
Line losses (MW)	10.8044	13.1376

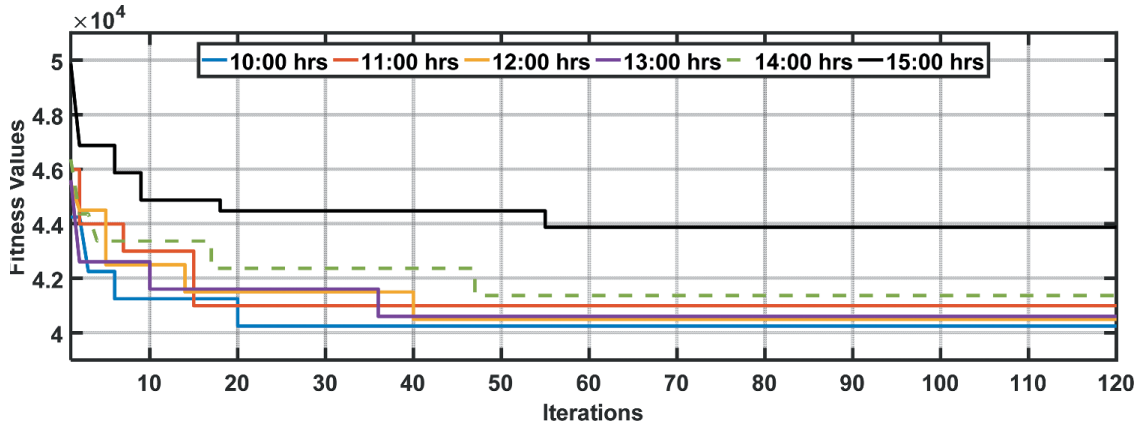


Figure 3. QI-TVIW-GSA-PSO-based convergence characteristics for SCEED at full solar radiation.

this region generally receives. Figure 5 shows the convergence characteristics of QI-TVIW-GSA-PSO for DCEED problem. It can be seen that the algorithm is effective for each case and converges rapidly.

6. Conclusion and scope of future work

In quantum space, there is no restriction in the movement of the object and it can move freely in the entire search space and can achieve optimality with even a reduced population. To exploit these

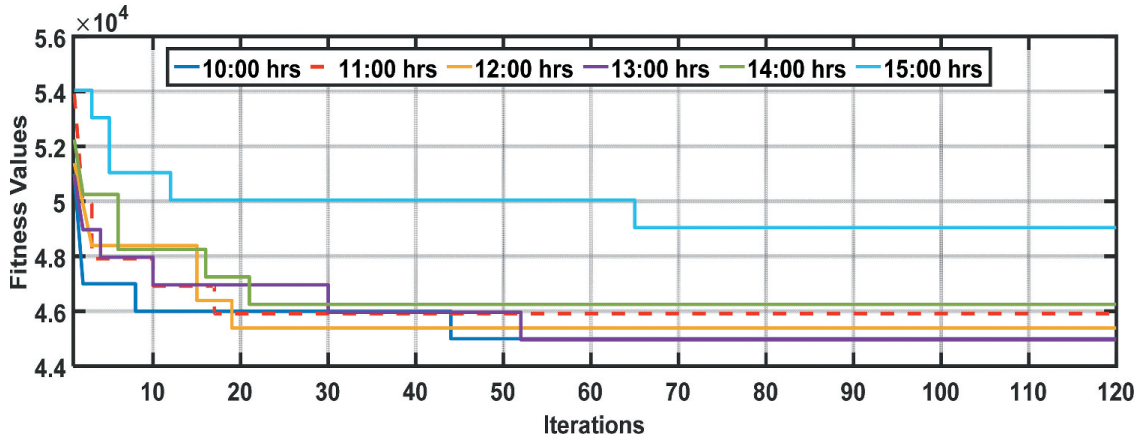


Figure 4. QI-TVIW-GSA-PSO-based convergence characteristics for SCEED at reduced solar radiation.

properties, the TVIW-GSA-PSO algorithm is applied through the quantum mechanics principle. An adaptive contraction expansion factor is also introduced in the position update equation of QI-TVIW-GSA-PSO for better exploration. A quantum binary TVIW-GSA-PSO algorithm is used to schedule solar panels in hybrid dispatch. These techniques are exploited to find the joint optimized solution.

CEED problem of a 10-unit 2000 MW load system considering valve point effect and transmission line losses is solved and compared to recent algorithms like NSGA-II, MODE, SPEA-II, FPA, PDE, MBA, and FPA for fuel cost and emission. QI-TVIW-GSA-PSO delivers the lowest fuel cost compared to other state-of-the-art algorithms leading to saving up to 1309 . USD The SCEED and DCEED of a hybrid solar thermal plant are solved using the proposed technique. The QI-TVIW-GSA-PSO optimizes the generating unit position using the quantum mechanics principle while the QB-TVIW-GSA-PSO uses the rotational function of quantum computing to predict the next position.

A 6-unit system is considered and the load demand for 24 h is taken from Tata Power-DDL power plant, India. A 13 solar PV plant along with thermal units are supplying the load and is considered to be operating in New Delhi, India. The SCEED problem is solved for full solar radiation and reduced solar radiation with line losses included. It can be seen from results that, with increased solar share, the fuel cost, the emission cost and the total cost reduces considerably. A reduction in solar share by 30% causes total cost to rise up to 12% and the emission is increased by 17%. In case of DCEED problem, the algorithm converges rapidly without violating any constraint. Simulation results demonstrate that, even with multiple constraints and limited search space, the proposed algorithm manages to optimize the problem efficiently and converges rapidly.

The proposed work utilizes the single-objective approach to solve the multiobjective function. The future work includes solving this MIOP using multiobjective approach. In addition, an option to turn off the generators in case of low power demand may be further investigated.

Algorithm 1. Steps of TVIW-GSA-PSO for mixed integer optimization CEED problem.

(1) Initialize population as $A_i = \left(a_i^1, \ldots a_i^d, \ldots a_i^n, u_i^1, \ldots u_i^d, \ldots u_i^n\right)$ consisting of both continuous and binary values.
(2) Evaluate the fitness function using (16).
(3) Update $\mathbb{G}(t)$, $M_i(t)$, Worst (t), finest (t) for i=1,2, ...,n.
(4) Calculate total force using (25).
(5) Calculate acceleration and inertia weight using (26) and (38), respectively.
(6) Update velocity and position of the objects using (37) and (39), respectively.
(7) Update the position (state) of the binary particles using (40).
(8) Repeat **Steps 2** to **7** until stopping criteria is met.

Table 8. DCEED for full day with reduced solar exposure.

Hour	P_1	P_2	P_3	P_4	P_5	P_6	P_{gen}	S/W	Fuel cost ($/h)	Emission (kg/h)	Emission cost ($/h)	Solar share	Solar cost	Total cost ($/h)
1.00	297.79	140.77	170.85	75.00	200.00	70.59	955.00		11209.00	965.77	8025.60			19234.60
2.00	284.79	140.77	170.85	75.00	190.00	80.59	942.00		11060.00	918.77	7345.49			18405.49
3.00	204.79	170.77	180.85	105.00	190.00	101.59	953.00		11313.00	823.55	7838.86			19151.86
4.00	194.79	170.77	170.85	102.00	190.00	101.59	930.00		11063.00	782.34	6998.74			18061.74
5.00	190.79	174.77	171.85	105.00	175.00	117.59	935.00		11138.00	776.50	5922.91			17060.91
6.00	200.79	160.77	189.85	111.00	196.00	104.59	963.00		11461.00	830.66	6857.85			18318.85
7.00	242.90	160.77	189.74	101.50	180.00	114.09	989.00		11676.00	899.06	7529.38			19205.38
8.00	244.90	180.77	199.74	141.40	150.00	106.19	1023.00		12071.00	970.46	7611.73			19682.73
9.00	259.89	196.77	229.74	150.00	170.41	119.19	1126.00		13343.42	1148.35	8584.98			21928.40
10.00	234.88	180.77	199.74	141.40	150.00	101.19	1007.99	A_1	11898.39	944.82	8192.56	142.02	28.36	20119.31
11.00	270.90	197.77	199.74	141.59	150.00	106.65	1066.65	A_2	12560.58	1068.23	9553.94	134.35	26.71	22141.23
12.00	236.55	196.77	229.74	150.00	170.41	119.19	1102.65	A_3	13098.14	1092.12	9278.09	132.35	27.02	22403.24
13.00	261.00	197.75	202.76	139.59	150.00	108.65	1059.75	A_4	12489.36	1047.74	7690.42	130.25	26.49	20206.28
14.00	297.01	199.77	175.74	140.23	180.24	119.19	1112.17	A_5	13150.97	1129.80	8148.20	138.83	28.24	21327.40
15.00	306.56	159.71	215.74	140.49	180.04	119.19	1121.73	A_6	13248.23	1168.83	8152.42	141.27	28.73	21429.39
16.00	359.89	196.77	253.70	150.00	170.47	119.17	1250.00		14778.33	1506.96	10447.50			25225.82
17.00	330.70	196.77	256.90	149.00	171.46	116.17	1221.00		14430.41	1420.15	10697.69			25128.10
18.00	310.70	197.77	256.90	149.00	171.46	116.17	1202.00		14212.18	1361.82	11534.99			25747.17
19.00	267.00	197.77	256.90	148.00	172.35	116.98	1159.00		13737.51	1241.62	10563.92			24301.43
20.00	238.34	160.77	256.90	148.21	171.14	116.64	1092.00		12978.68	1102.58	8180.71			21159.40
21.00	234.76	190.90	199.41	141.74	149.00	107.19	1023.00		12087.86	965.23	7418.38			19506.24
22.00	232.90	169.77	189.74	122.50	160.09	109.00	984.00		11617.03	888.36	7304.30			18921.33
23.00	223.77	177.90	180.74	127.50	158.09	107.00	975.00		11529.46	867.50	7758.85			19288.30
24.00	223.77	178.90	181.74	109.00	158.59	108.00	960.00		11328.78	852.34	7403.20			18731.99

A_1 = [0 1 1 0 1 1 1 0 1 0 1 1 1], A_2 = [1 1 0 1 1 0 1 0 1 1 1 1 0], A_3 = [1 1 0 0 0 1 1 1 1 1 0 1 1], A_4 = [1 1 0 1 0 1 0 1 1 1 0 1 1], A_5 = [1 1 0 1 0 1 0 1 1 1 0 1 1], A_6 = [1 1 1 1 0 1 0 1 1 0 1 1 0].

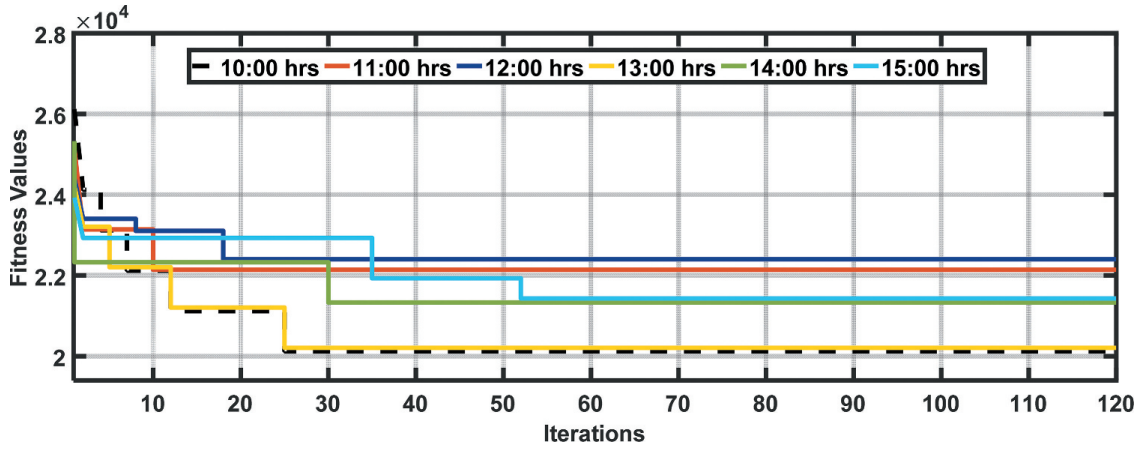


Figure 5. QI-TVIW-GSA-PSO-based convergence characteristics for DCEED at reduced solar radiation.

Algorithm 2. Steps of QI- TVIW-GSA-PSO for mixed integer-based CEED problem.

(1) Initialize population as $A_i = \left(a_i^1, \ldots a_i^d, \ldots a_i^n, u_i^1, \ldots u_i^d, \ldots u_i^n\right)$ consisting of both continuous and binary values.
(2) Evaluate the fitness function using (16).
(3) Update $\mathbb{G}(t)$, $M_i(t)$, Worst (t), finest (t) for i=1,2, ...,n.
(4) Evaluate total force using (25).
(5) Calculate acceleration and inertia weight using (26) and (38), respectively.
(6) Calculate velocity and position of the objects using (37) and (39), respectively.
(7) Update the position (state) of the binary particles using (40).
(8) Calculate the fitness function and evaluate best position.
(9) Update the position using the quantum mechanics equation (48) for thermal generating units.
(10) Update the switching state of solar unit using (56).
(11) Repeat **Steps 2 –10** until stopping criteria is met.

List of abbreviations

CEED Combined EEDCLQPSO Quantum-behaved PSO with collaborative attractors DCEED Dynamic CEEDDE Differential evolution EAE volutionary algorithm EED Economic emission dispatch ELD Economic load dispatch EP Evolutionary programing FPA Flower pollination algorithm GA Genetic algorithm GP Goal programming GSA Gravitational search algorithm HSP Hybrid sequential programming LCOE Levalized cost of electricity LMNN Levenberg Marquardt neural network MBA Mine blasting algorithm MIOP Mixed integer optimization problem MODE Multi-objective DENSGA-II Non-dominated sorting GA-IIPDE Pareto DEPOZ Prohibited operating zone PSO Particle swarm optimization QBA Quantum behaved bat algorithm QI-TVIW-GSA-PSO Quantum inspired time varying inertia weight gravitational search algorithm PSORRL Ramp rate limit SCEED Static CEEDSPEA-2 Strength Pareto EA 2TLBO Teaching-learning-based optimization

Disclosure statement

The authors declare that they have no conflict of interest.

ORCID

Kapil Deo Bodha http://orcid.org/0000-0003-2862-2170

List of symbols

$\left\|alpha_{i,d}^{t+1}\right\|^2$ and $\left\|beta_{i,d}^{t+1}\right\|^2$	Probability of the position of $u_{i,d}^{t}$ in "0" state and "1" state, respectively
$Acl_i^d(t)$	Acceleration of the *i*th object at time *t* in the *d*th dimension
A_i,B_i,C_i,D_i and E_i	Fuel cost coefficients of the *i*th generating unit
$B_{i,j}$	Loss coefficient matrix
$F_i^d(t)$	Total force acting upon the *i*th object in the *d*th dimension
F_{ij}^d(t)	Force acting on the *i*th mass on the *j*th mass at time *t*
K_1 and K_2	Constants with values lying in [0, 1]
$M_{activej}(t)$	Active gravitational mass of the *j*th object at time *t*
$M_{inertiali}$(t)	Inertial mass of the *i*th object
$M_{passivei}(t)$	Passive gravitational mass of the *i*th object at time *t*
P_1 and P_2	Random number between [0, 1]
P_{rated}	Rated power of solar panel
P_s	Power generated by slack generator
P_{solar}	Power generated by a solar plant
RU_i and RD_i	Up rate and down rate of the *i*th generating unit, respectively
SR_i	Incident solar radiation
SW_j	Switching state of the *j*th solar plant
$Vel_i^d(t)$	Velocity of the *i*th object at time *t*
a_i^d	Position of the *i*th object at the *d*th dimension
$a_i^d(t)$	Position of the *i*th object at time *t*
$alpha_i$,$beta_i$,$gamma_i$,L_i and Z_i	Emission coefficients of the *i*th generating unit
$d_{ij}(t)$	Euclidian distance between two *i*th and the *j*th objects
d_{ij}^h	Hamming distance between active mass and passive mass
$t_{Ambient}$	Ambient temperature
B	Loss coefficient matrix
C	Objective function
EC_i	Emissions in kg/h of the *i*th generating unit
FC_i	Fuel cost of the *i*th generating unit
$G(t)$	Gravitational constant at time *t*
Globest	Global best solution
h	Price penalty factor
$\hat{H}$	Hamiltonian operator
M	Number of solar plants
N	Total number of generating units
P_{demand}	Power demand
P_{gi}	Power generated by the *i*th unit
$P_{igenmax}$	Minimum power limit of the *i*th generating unit
$P_{igenmin}$	Minimum power limit of the *i*th generating unit
P_{loss}	Transmission line loss
R()	Rotation transformation function
rand	Random number in [0, 1]
Δ	Rotation factor
σ	Global inertia weight
ϕ	Rotation angle
$\mathbb{G}(t_0)$	Gravitational constant at the first cosmic quantum interval at time t_0
$\mathbb{G}(t)$	Gravitational constant
$\Psi(f,t)$	Wave function of an object in quantum space expressed in terms of position (x) and time$t$
τ	Temperature coefficient

References

Abdelaziz, A. Y., E. S. Ali, and S. M. Abd Elazim. 2016. Combined economic and emission dispatch solution using flower pollination algorithm. *International Journal of Electrical Power and Energy Systems* 80:264–74. doi:10.1016/j.ijepes.2015.11.093.

Ali, E. S., and S. M. Abd Elazim. 2016. Mine Blast Algorithm for environmental economic load dispatch with valve loading effect. *Neural Computing & Applications* 30 (1):261–70. doi:10.1007/s00521-016-2650-8.

Basu, M. 2011. Economic environmental dispatch using Multi-Objective differential evolution. *Applied Soft Computing Journal* 11 (2):2845–53. https://doi.org/10.1016/j.asoc.2010.11.014 .

Benasla, L., A. Belmadani, and M. Rahli. 2014. Spiral optimization algorithm for solving combined economic and emission dispatch. *International Journal of Electrical Power and Energy Systems* 62:163–74. doi:10.1016/j.ijepes.2014.04.037.

Bodha, K. D., V. K. Yadav, and V. Mukherjee. 2020. Formulation and application of quantum-inspired tidal firefly technique for multiple-objective mixed cost-effective emission dispatch. *Neural Computing & Applications* 32 (13):9217–32. doi:10.1007/s00521-019-04433-0.

Bombay, IIT. n.d. "Solar thermal software simulator." https://www.ese.iitb.ac.in/oldweb/simulator/simulator.html .

Bulut, U., and A. Menegaki. 2020. Solar Energy-Economic growth nexus in Top 10 countries with the highest installed capacity. *Energy Sources, Part B: Economics, Planning and Policy* 1–14. doi:10.1080/15567249.2020.1788192.

Centre, delhi state load dispatch. n.d. "No Title." https://www.delhisldc.org/Redirect.aspx?Loc=0805 .

Chatterjee, A., S. P. Ghoshal, and V. Mukherjee. 2012. Solution of combined economic and emission dispatch problems of power systems by an opposition-Based harmony search algorithm. *International Journal of Electrical Power and Energy Systems* 39 (1):9–20. doi:10.1016/j.ijepes.2011.12.004.

Chaturvedi, K. T., M. Pandit, and L. Srivastava. 2006. "On-Line solution to combined economic and emission dispatch problem." In *Proceedings of the IEEE International Conference on Industrial Technology*. Mumbai, India. https://doi.org/10.1109/ICIT.2006.372451 .

Dhillon, J. S., S. C. Parti, and D. P. Kothari. 1993. Stochastic economic emission load dispatch. *Electric Power Systems Research* 26 (3):179–86. doi:10.1016/0378-7796(93)90011-3.

Duman, S., Y. Sönmez, U. Güvenç, and N. Yörükeren. 2012. Optimal reactive power dispatch using a gravitational search algorithm. *IET Generation, Transmission and Distribution* 6 (6):563. doi:10.1049/iet-gtd.2011.0681.

Elaiw, A. M., X. Xia, and A. M. Shehata. 2012. Application of model predictive control to optimal dynamic dispatch of generation with emission limitations. *Electric Power Systems Research* 84 (1):31–44. doi:10.1016/j.epsr.2011.09.024.

Elanchezhian, E. B., S. Subramanian, and S. Ganesan. 2016. Economic dispatch of energy sources with restrictions on fuel using TLBO algorithm. *Energy Sources, Part B: Economics, Planning and Policy* 11 (10):920–28. doi:10.1080/15567249.2013.792398.

Fan, J. Y., and L. Zhang. 1998. Real-Time economic dispatch with line flow and emission constraints using quadratic programming. *IEEE Transactions on Power Systems*. doi:10.1109/59.667345.

Gaing, Z. L. 2003. Particle swarm optimization to solving the economic dispatch considering the generator constraints. *IEEE Transactions on Power Systems* 18 (3):1187–95. doi:10.1109/TPWRS.2003.814889.

Government of India,Central Eectricity Authority, Ministry of power. 2020. All India Installed Capacity (in Mw) of power stations installed capacity (in Mw) of POWER UTILITIes in the States/Uts located in. *Central Eectricity Authority, Ministry of Power* 4:1–7. https://doi.org/http://www.cea.nic.in/reports/monthly/installedcapacity2016/installed_capacity-03.pdf .

Güvenç, U., Y. Sönmez, S. Duman, and N. Yörükeren. 2012. combined economic and emission dispatch solution using gravitational search algorithm. *Scientia Iranica* 19 (6):1754–62. doi:10.1016/j.scient.2012.02.030.

India, Ministry of Power Govt. of. 2020. "Power Sector at a Glance ALL INDIA." 18 May. 2020. Ministry of Power Govt. of India.

Ji, B., X. Yuan, L. Xianshan, Y. Huang, and L. Wenwu. 2014. Application of quantum-inspired binary gravitational search algorithm for thermal unit commitment with wind power integration. *Energy Conversion and Management* 87:589–98. doi:10.1016/j.enconman.2014.07.060.

Khan, N. A., A. B. Awan, A. Mahmood, S. Razzaq, A. Zafar, and G. A. S. Sidhu. 2015. Combined emission economic dispatch of power system including solar photo voltaic generation. *Energy Conversion and Management* 92:82–91. doi:10.1016/j.enconman.2014.12.029.

Khan, N. A., G. A. S. Sidhu, and F. Gao. 2016. Optimizing combined emission economic dispatch for solar integrated power systems. *IEEE Access* 1–1. doi:10.1109/ACCESS.2016.2587665.

King, R. T. F. A., and H. C. S. Rughooputh. 2003. "Elitist multiobjective evolutionary algorithm for environmental/economic dispatch." In *2003 Congress on Evolutionary Computation, CEC 2003 - Proceedings*. Canberra, ACT, Australia. https://doi.org/10.1109/CEC.2003.1299792 .

Koltsaklis, N. E., A. S. Dagoumas, and V. Mladenov. 2020. Electricity market clearing algorithms: A case study of the bulgarian power system. *Energy Sources, Part B: Economics, Planning and Policy*. doi:10.1080/15567249.2020.1845252.

Li, Y., J. Wang, D. Zhao, L. Guoqing, and C. Chen. 2018. A Two-Stage approach for combined heat and power economic emission dispatch: Combining Multi-Objective optimization with integrated decision making. *Energy* 162:237–54. doi:10.1016/j.energy.2018.07.200.

Liang, H., Y. Liu, L. Fengzhong, and Y. Shen. 2018. A multiobjective hybrid bat algorithm for combined economic/Emission dispatch. *International Journal of Electrical Power and Energy Systems* 101 (March):103–15. doi:10.1016/j.ijepes.2018.03.019.

Liu, Z. 2018. What is the future of solar energy? Economic and policy barriers. *Energy Sources, Part B: Economics, Planning and Policy* 13 (3):169–72. doi:10.1080/15567249.2017.1416704.

Mahdi, F. P., M. Pandian Vasant, V. Kallimani, J. Watada, and J. Watada; . 2019. Quantum-Behaved Bat algorithm for many-objective combined economic emission dispatch problem using cubic criterion function. *Neural Computing & Applications* 31 (10):5857–69. doi:10.1007/s00521-018-3399-z.

Mahdi, F. P., P. Vasant, M. Abdullah-Al-Wadud, J. Watada, and V. Kallimani. 2018. A Quantum-Inspired particle swarm optimization approach for environmental/Economic power dispatch problem using cubic criterion function. *International Transactions on Electrical Energy Systems* 28 (3):e2497. doi:10.1002/etep.2497.

Manteaw, E. D., and N. A. Odero. 2012. Multi-Objective Environmental/Economic dispatch solution using ABC_PSO Hybrid Algorithm. *International Journal of Scientific and Research Publications 2*(12), 1-7.

Meng, K., H. G. Wang, Z. Y. Dong, and K. P. Wong. 2010. Quantum-Inspired particle swarm optimization for valve-point economic load dispatch. *IEEE Transactions on Power Systems* 25 (1):215–22. doi:10.1109/TPWRS.2009.2030359.

Mirjalili, S., and S. Z. M. Hashim. 2010. "A New Hybrid PSOGSA algorithm for function optimization." In *Proceedings of ICCIA 2010-2010 International Conference on Computer and Information Application*. Tianjin, China. https://doi.org/10.1109/ICCIA.2010.6141614 .

Nations, United. 2015. "The Paris agreement." paris. https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement .

Perkins, G. 2020. Perspectives and Economics of combining biomass liquefaction with solar PV for energy storage and electricity production. *Energy Sources, Part B: Economics, Planning and Policy* 1–17. doi:10.1080/15567249.2020.1749910.

Ramachandra, T. V., R. Jain, and G. Krishnadas. 2011. Hotspots of solar potential in India. *Renewable and Sustainable Energy Reviews* 15 (6):3178–86. doi:10.1016/j.rser.2011.04.007.

Rashedi, E., H. Nezamabadi-pour, and S. Saryazdi. 2009. GSA: A Gravitational Search Algorithm. *Information Sciences* 179 (13):2232–48. doi:10.1016/j.ins.2009.03.004.

Rashedi, E., H. Nezamabadi-Pour, and S. Saryazdi. 2010. BGSA: Binary Gravitational Search Algorithm. *Natural Computing* 9 (3):727–45. doi:10.1007/s11047-009-9175-3.

Ravikumar, P. V., B. K. Panigrahi, W. C. Hong, and R. Sharma. 2014. A multiobjective bacterial foraging algorithm to solve the environmental economic dispatch problem. *Energy Sources, Part B: Economics, Planning and Policy* 9 (3):236–47. doi:10.1080/15567249.2010.485167.

REN21. 2018. Renewables 2018 global status report (Paris: REN21 secretariat). *Paris: Renewable Energy Policy Network for the 21st Century*. 978-3-9818911-3-3.

Ruhe, C. H. W. 2019. Statistical Review. *JAMA: The Journal of the American Medical Association* 225 (3):299–306. doi:10.1001/jama.1973.03220300055017.

Sakthivel, V. P., M. Suman, and P. D. Sathya. 2020. Squirrel search algorithm for economic dispatch with Valve-Point effects and multiple fuels. *Energy Sources, Part B: Economics, Planning and Policy* 15 (6):351–82. doi:10.1080/15567249.2020.1803451.

Shaw, B., V. Mukherjee, and S. P. Ghoshal. 2012. A Novel Opposition-Based gravitational search algorithm for combined economic and emission dispatch problems of power systems. *International Journal of Electrical Power and Energy Systems* 35 (1):21–33. doi:10.1016/j.ijepes.2011.08.012.

Shen, X., D. Zou, N. Duan, and Q. Zhang. 2019. An Ef Fi Cient Fi Tness-based differential evolution algorithm and a constraint handling technique for dynamic economic emission dispatch. *Energy* 186:115801. doi:10.1016/j.energy.2019.07.131.

Shilaja, C., and K. Ravi. 2017. Optimization of Emission/Economic dispatch using euclidean affine flower pollination algorithm (EFPA) and Binary FPA (BFPA) in Solar Photo voltaic generation. *Renewable Energy* 107:550–66. doi:10.1016/j.renene.2017.02.021.

Stavrakas, V., S. Papadelis, and A. Flamos. 2019. An agent-based model to simulate technology adoption quantifying behavioural uncertainty of consumers. *Applied Energy* 255 (May):113795. doi:10.1016/j.apenergy.2019.113795.

Sundaram, A. 2018. Solution of combined economic emission dispatch problem with Valve-Point effect using hybrid NSGA II-MOPSO. *Particle Swarm Optimization with Applications*. doi:10.5772/intechopen.72807.

Sundaram, A. 2020. Multiobjective Multi-Verse optimization algorithm to solve combined economic, heat and power emission dispatch problems. *Applied Soft Computing Journal* 91:106195. doi:10.1016/j.asoc.2020.106195.

Wang, G., Y. Zha, W. Ting, J. Qiu, J.-C. Peng, and X. Gang. 2020. Cross entropy optimization based on decomposition for Multi- Objective economic emission dispatch considering renewable energy generation uncertainties. *Energy* 193:116790. doi:10.1016/j.energy.2019.116790.

Xin-gang, Z., J. Liang, M. Jin, and Z. Ying. 2020a. An improved quantum particle swarm optimization algorithm for environmental economic dispatch. *Expert Systems with Applications* 152:113370. doi:10.1016/j.eswa.2020.113370.

Xin-gang, Z., Z. Ze-qi, X. Yi-min, and M. Jin. 2020b. Economic-Environmental dispatch of microgrid based on improved quantum particle swarm optimization. *Energy* 195:117014. Elsevier Ltd. doi:10.1016/j.energy.2020.117014.

Xue, H., Y. Bai, H. Hongping, X. Ting, and H. Liang. 2019. A novel hybrid model based on TVIW-PSO-GSA algorithm and support vector machine for classification problems. *IEEE Access* 7:27789–801. doi:10.1109/ACCESS.2019.2897644.

Yazdani, S., H. Nezamabadi-Pour, and S. Kamyab. 2014. A gravitational search algorithm for multimodal optimization. *Swarm and Evolutionary Computation* 14:1–14. doi:10.1016/j.swevo.2013.08.001.

Zhang, Y., L. Tie, N. Guangyu, L. Guoqing, and L. Yang. 2015. Optimized extreme learning machine for power system transient stability prediction using synchrophasors. *Mathematical Problems in Engineering* 2015:1–8. doi:10.1155/2015/529724.

Appendix

Table A1. Fuel cost coefficient and generation constraints of IEEE-6 generator system.

Generator	a ($/MW^2 h)	b ($/MW h)	c ($/h)	P_{min} (MW)	P_{max} (MW)
1	0.152	38.539	756.798	10	125
2	0.105	46.159	451.325	10	150
3	0.028	40.396	1049.325	40	250
4	0.035	38.305	1243.531	35	210
5	0.021	36.327	1658.569	130	325
6	0.017	38.270	1356.270	125	315

Table A2. Coefficient of emission.

Generator	α (kg/MW^2 h)	β (kg/MW h)	Υ (kg/h)
1	0.00419	0.32767	13.85932
2	0.00419	0.32767	13.85932
3	0.00683	−0.54551	40.26690
4	0.00683	−0.54551	40.26690
5	0.00461	−0.51116	42.89553
6	0.00461	−0.51116	42.89553

Table A3. Solar plant power rating and per unit rate.

Solar plant	P_{rated} (MW)	Unit rate ($/kW h)	Solar plant	P_{rated} (MW)	Unit rate ($/kW h)
1	20	0.15	8	40	0.22
2	25	0.16	9	40	0.22
3	25	0.16	10	40	0.22
4	30	0.18	11	40	0.22
5	30	0.18	12	40	0.22
6	35	0.19	13	40	0.22
7	35	0.19			

Table A4. Solar radiation, temperature and power demand for New Delhi, India on 3 May 2020.

Time (h)	Temperature (°C)	Power demand (MW)	Solar radiation (W/m^2)	Time (h)	Temperature (°C)	Power demand (MW)	Solar radiation (W/m^2)
1:00	30.4	945.23	0	13:00	38	698.62	444.6
2:00	29.7	905.83	0	14:00	38.7	719.57	471.4
3:00	29	858.1	0	15:00	39.1	753.37	479.7
4:00	28.4	828.23	0	16:00	39.1	759.01	457.5
5:00	27.8	807.58	0	17:00	38.8	724.09	404.4
6:00	27.3	761.21	222.4	18:00	38.3	715.98	323.9
7:00	27.5	735.08	323.9	19:00	36.9	722.88	0
8:00	28.9	720.55	404.4	20:00	35.3	803.35	0
9:00	31.6	694.94	457.5	21:00	34.1	862.86	0
10:00	33.7	699.95	479.7	22:00	33	911.59	0
11:00	35.5	705.73	471.4	23:00	32	956.49	0
12:00	37	697.58	444.6	00:00	31	992.37	0

Table A5. Fuel cost coefficient, emission coefficient and ramp rate constraints.

Gen	a ($/MW^2 h)	b ($/MW h)	c ($/h)	α (kg/MW^2 h)	β (kg/MW h)	ϒ (kg/h)	P_{min} (MW)	P_{max} (MW)	UR (MW/h)	DR (MW/h)
1	0.007	7.0	240	0.00419	0.32767	13.8593	100	500	80	120
2	0.009	10.0	200	0.00419	0.32767	13.8593	50	200	50	90
3	0.009	8.0	220	0.00683	−0.54551	40.2669	80	300	65	100
4	0.009	11.0	200	0.00683	−0.54551	40.2669	50	150	50	90
5	0.008	10.5	220	0.00461	−0.51116	42.8955	50	200	50	90
6	0.007	12.0	190	0.00461	−0.51116	42.8955	50	120	50	90

Table A6. Load demand for a day.

Hour	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Demand (MW)	955	942	953	930	935	963	989	1023	1126	1150	1201	1235
Hour	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
Demand (MW)	1190	1251	1263	1250	1221	1202	1159	1092	1023	984	975	960

Sensors and Actuators A 331 (2021) 112926

Contents lists available at ScienceDirect

Sensors and Actuators A: Physical

journal homepage: www.elsevier.com/locate/sna

A simple and novel SAW magnetic sensor with PVA bound magnetostrictive nanopowder film

Fahim [a], Mainuddin [b,*], Pooja Rajput [a], Jitender Kumar [a], A.T. Nimal [a]

[a] Solid State Physics Laboratory, Delhi, 54, India
[b] Department of Electronics and Communication Engineering, Jamia Millia Islamia University, Delhi, 25, India

ARTICLE INFO

Article history:
Received 2 March 2021
Received in revised form 2 June 2021
Accepted 15 June 2021
Available online 18 June 2021

Keywords:
Magnetic field sensor
Surface Acoustic Wave
Magnetostriction
Polyvinyl Alcohol

ABSTRACT

- In this paper a Surface Acoustic Wave (SAW) magnetic field sensor based on magnetostriction effect was realized using a simple, inexpensive and novel approach. A single port SAW resonator on ST-Quartz substrate was used and configured as a Colpitt Oscillator. A novel coating approach was adopted whereby a thin layer of Polyvinyl Alcohol (PVA) was coated on bare SAW device to passivate and electrically isolate the Inter Digital Transducers (IDTs) followed by thin layer of PVA nanocomposite magnetostrictive coating applied using drop dry method. Separate sensors were realized using Nickel (Ni) and Iron (Fe) materials and their performances were evaluated using a Helmholtz coil. Good magnetic sensitivity and repeatability was observed in developed sensors. Excellent long term stability was achieved and no fatigue was observed.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

Magnetic sensors find many applications in industrial, medical, automotive and scientific fields [1]. From contactless current sensing, linear and angular position to rotation sensing, these sensors have distinct advantages over optical or capacitive sensors in dusty, humid, oily and greasy environments [2]. Their use became widespread particularly after solid-state sensor devices became available. In recent times, many consumer products like cameras, printers and scanners have come up as a big market for magnetic sensors. Other rapidly growing application of magnetic sensors are gaming consoles and smart mobile phones where compass has become an essential feature [3].

The most commonly used technologies for sensing magnetic fields are Hall Effect [4], magneto-resistance [5], magneto-diode [6], magneto-transistor [7], MEMS based sensors [8], optical [9], giant magneto-impedance (GMI) [10], giant magnetoresistance (GMR) [11], superconducting quantum interference device (SQUID) [12], fluxgate [13], coil and many more [14]. The reason for such diversity is the wide range of applications.

Surface Acoustic Wave (SAW) based magnetic field sensor are another type of solid-state magnetic sensors. Although not fully exploited, a unique advantage it offers over other technologies is its ability to work in remote and powerless (passive) manner [15–17]. SAW devices are small, low cost, rugged, mass producible, very sensitive and linear in a large range [18,19]. SAW magnetic sensor can be realized by depositing a thin magnetostrictive material layer on SAW delay line [20] or by fabricating SAW interdigital transducers (IDT) by magnetostrictive material itself [21]. SAW can also be combined with a GMI microwire to form a SAW-GMI magnetic sensor [22]. While the latter two methods involve complex designing and fabrication of SAW device, the first method uses a conventional and commercially available SAW device. The deposition of magnetostrictive coating is achieved by dc magnetron sputtering, vapour deposition using PVD/CVD, Electroplating, electro spray etc. These techniques are, however, complex, require vacuum conditions, well controller atmosphere and use costly and energy intensive equipment [23]. Additionally, these films also suffer from significant mechanical fatigue and deterioration of performance over a period of operation [24]. Further, for SAW devices like resonators and SAW tags, an additional passivation layer is also required to electrically insolate IDTs from the metal coating.

In this paper, a simple and low cost SAW magnetic sensor was developed by coating magnetostrictive nanocomposite through optimized and advanced drop dry method. Nanocomposite mate-

* Corresponding author.
E-mail addresses: fahimdurani1@gmail.com (Fahim), mainuddin@jmi.ac.in (Mainuddin), poojarajput0127@gmail.com (P. Rajput), jitender.tilwaliya@gmail.com (J. Kumar), atnimal@yahoo.com (A.T. Nimal).

https://doi.org/10.1016/j.sna.2021.112926
0924-4247/© 2021 Elsevier B.V. All rights reserved.

rials consist of magnetic nano particles with shape anisotropy embedded in an elastomer matrix [25]. In this study we explored PVA as binding material to bind the ferromagnetic metal particles within thin film and also as passivation layer. When magnetic field is applied, it forces molecular dipoles of ferromagnetic material to align in the direction of the field. The forced alignment causes it to deform mechanically (shrink or stretch) producing strain known as magnetostrictive effect. As the metal particles are embedded in PVA, the magnetostrictive strain is translated onto the device surface disturbing the SAW propagation as it is very sensitive to changes on the surface of substrate. The experiments were conducted using nanopowders of well- known and thoroughly researched ferromagnetic materials Fe and Ni having positive and negative magnetostrictive coefficients respectively. The coated devices were hermetically sealed to avoid the effects of humidity on PVA. The performance of the developed sensors was characterized experimentally by using Helmholtz Coil. Good sensitivity and repeatability were observed. The main features of our novel approach over the existing techniques are as follows:

- An alternative and first ever PVA bound nano-composite magnetostrictive SAW detector.
- Simple and low-cost method. Doesn't need costly coating equipment.
- Immune to fatigue related film damage and rupture.

2. Saw Magnetic Sensor

2.1. SAW sensor device

Commercial SAW resonator device chip (Murata RO3101) having frequency of 433.92 MHz were used. These one-port resonators are fabricated on quartz substrate with aluminium IDTs having ~7-micron spatial period. The acoustically reflective gratings are located at both ends of substrate external to IDTs. The device has unloaded Q factor of 7400. Two similar SAW resonators, one with magnetostrictive coating for sensing and second without coating as reference, were used for experiment. Both resonator chips were separately mounted on a PCB header as shown in Fig. 1(a) and sealed hermetically. Electrically the resonators are connected as feedback elements of Colpitt oscillators (Nimal et al. [30]) and used in differential mode by mixing their outputs. Fig. 1(b) shows the schematic of our SAW magnetic sensor. A difference frequency (few kHz) exists between sensor and reference devices due to the mass loading in sensor device because of sensor coating. When the sensor device was exposed to a magnetic field, the frequency changed w.r.t reference device in proportion to the field strength. This difference frequency was measured using our patented auto ranging reciprocal frequency counter implemented in FPGA [31]. The measured difference frequency was stored and plotted in a PC in real time. The frequency stability was measured for both long term and short term and typical values obtained are ±15 Hz and ±5 Hz respectively as shown in Fig. 2.

2.2. Magnetostrictive coating

The magnetostrictive coating forms the most important aspect of our research. Both electrical and mechanical properties of coating material affect SAW propagation. The amount of mass to be loaded and avoiding the finger shorting are among the critical factors to be taken care off. Iron and nickel are among the most well known ferromagnetic materials [32]. The magnetostriction mechanism of Fe and Ni has been studied extensively by researchers and numerous publications have been published. Magnetic anisotropy

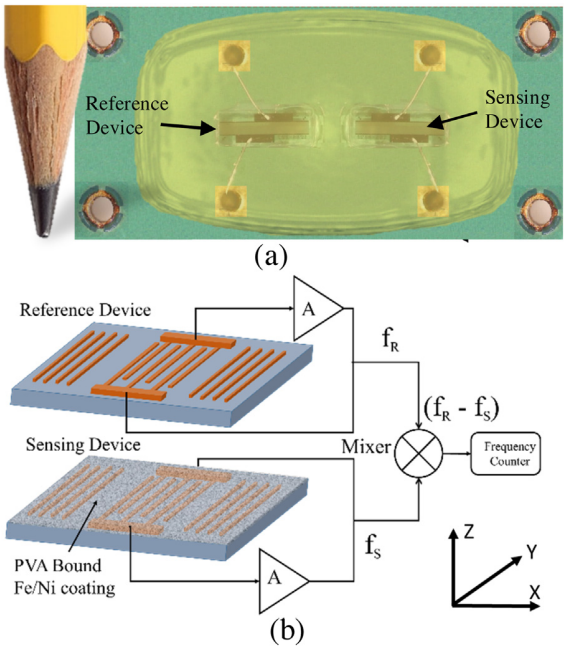


Fig. 1. (a) SAW resonator mounting arrangement (b) Scheme and working principle of differential SAW magnetic field sensor.

also has a crucial role as the magnetostrictive coefficient and its sign (in case of Fe) are dependent on it. The particle dimension is another significant factor. Magnetic nano-dimensional particles possess unique physical and chemical properties in comparison to bigger sized particles [33]. Binding of metal particles on the crystal surface is also an important parameter. The particles should be bound strongly enough to move synchronously with acoustic wave displacement and withstand magnetic force without being knocked out. For this study we used PVA as passivation and binding material.

PVA is a semi-crystalline and non-toxic polymer with excellent chemical stability [34]. It is used in diverse areas like coating, binding, membrane etc. due to its outstanding mechanical and film forming properties [35]. PVA is hydrophilic and hence has widely been used on SAW for humidity sensing. It, however, has excellent binding properties and in the present study we have exploited this property of PVA in a novel way to realize SAW magnetic sensor. First a thin layer of PVA was coated on the device basically to cover the whole device (passivation) and then a second layer of PVA with metal nanopowder was drop cast on it. The advantages of polymer based magnetic nanoparticle films are manifold. Polymers are known to be excellent host matrices for metallic nanoparticles [36]. It traps the nanoparticles and restricts the agglomeration thereby improves the chemical stability [37]. It is light weight, corrosion resistant and has high mechanical strength [38]. It can easily transform into any shape and hence well suited for SAW device with electrodes on the surface.

For passivation, PVA solution was prepared by dissolving 1 mg PVA in 10 ml de-ionized water with constant stirring for 3 h. Before coating, the device was subjected to oxygen plasma cleaning in order to remove any residual impurity. The PVA solution was then spin coated on SAW device using a drop of 5 μl solution and spinning at 4500 rpm for 30 s followed by curing at 80 °C for 4 h. The film uniformly covered the whole device including IDTs and reflectors. 3D Atomic force microscopy (AFM) characterization was conducted to examine the surface structure of thin PVA coating and the morphology shown in Fig. 3(a) indicates a flat and uniform surface.

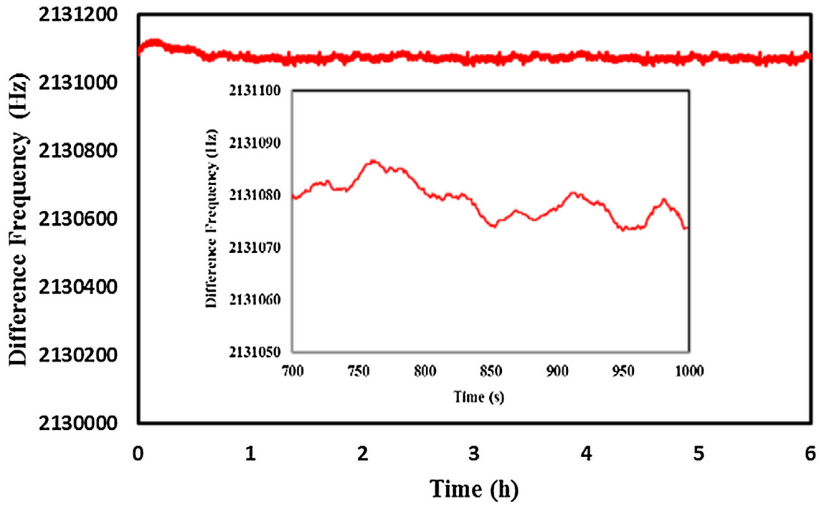


Fig. 2. Frequency stability test results.

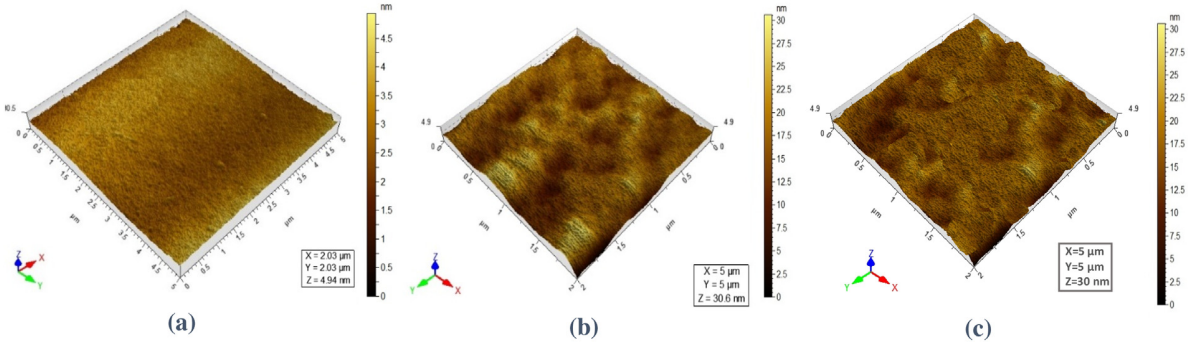


Fig. 3. AFM images of (a) PVA coated for passivating SAW IDTs (b) Ni coated SAW device (c) Fe coated SAW device.

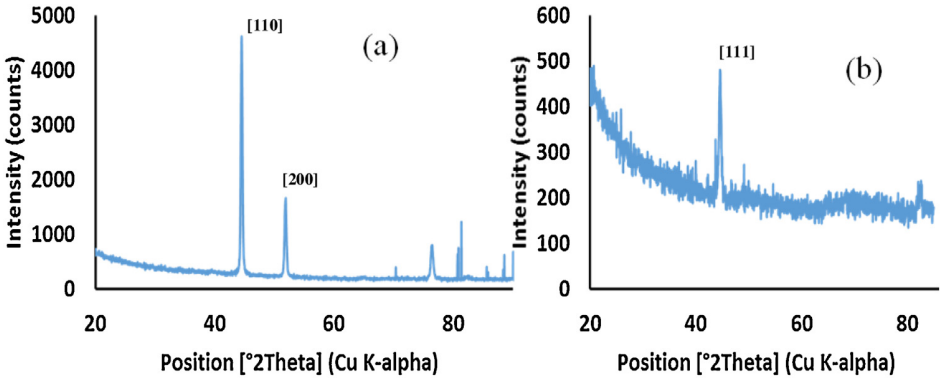


Fig. 4. X-ray diffraction pattern of fabricated SAW magnetic sensors (a) Fe coated (b) Ni Coated.

In order to prepare the ferromagnetic metal nano particle solution with binder, we used 1 mg Nano powder of Iron/Nickle (Sigma Aldrich) and mixed in 1 ml of PVA solvent. The solution of PVA with metal particles was ultra-sonicated for four hours at 50 °C. It was coated on the surface of the SAW device by drop dry method using 1 μl solution. We used a drop coating system which consists of a MicroDot valve (Model 741 MD-SS) and dispensing robot (model DR2203, Make Nordson). MicroDot valve is a pneumatically operated adjustable needle valve and designed to apply very precise deposits down to fractions of a microlitre. The coated devices were kept at 80 °C for four hours in an oven. Fig. 3(b) and (c) show the 3D AFM image of coated Ni and Fe films indicating well formed, uniform and flat morphologies. The crystallographic quality of the coated samples was also checked by x-ray diffraction (XRD) method as shown in Fig. 4(a) & (b). The crystalline positions of both samples were in conformity to known orientations. In order to achieve suitable coating level which is optimum in terms of mass load and sensitivity to magnetic field, an iterative study was conducted where component of Ni was varied from 1 to 5 mg in 1 ml of PVA solution and response to magnetic field was observed as shown in Fig. 5. The response obtained showed that as we increase the concentration of metal powder on the device, the

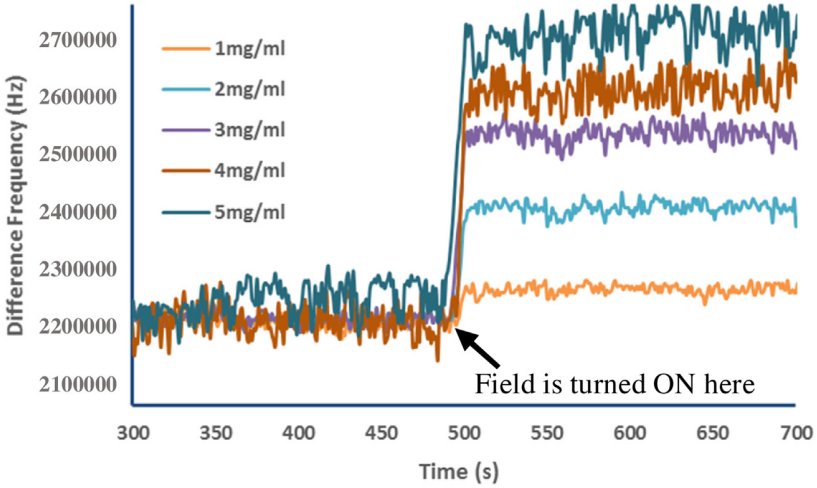


Fig. 5. Normalized frequency response at five concentration levels of Ni deposited on SAW device. 120 m T was applied.

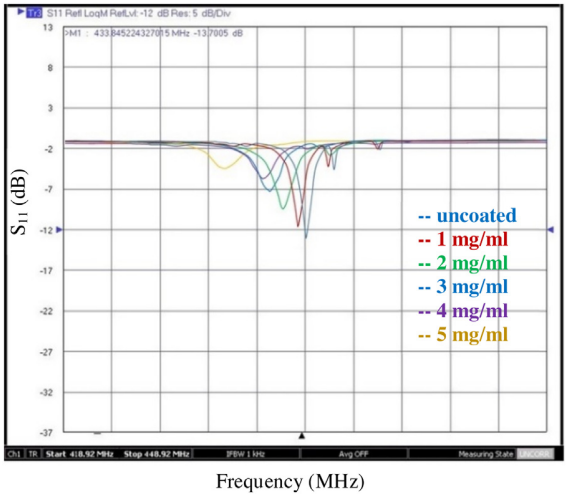


Fig. 6. S_{11} Plot of uncoated and coated devices at different nanopowder density levels.

sensitivity increased. The findings are in conformity with literature [39].

In order to study the performance at different coating levels, the devices were characterized using VNA (Model MS4642A, Make Anritsu) and insertion loss was recorded for both uncoated and coated devices. Fig. 6 shows S_{11} plot of devices coated with different quantities of metal particles in 1 ml of PVA solution. Table 2 presents the characterization data for different coating parameters. Both Fig. 6 and Table 2 indicate that as we increase the quantity of metal particles the device frequency decreases due to mass load on the substrate. It is further accompanied by decrease in quality-factor as indicated by broadening of curves and increase in insertion loss going upto −7.91 dB for maximum coating. With higher coating, the average noise also increased substantially and device stability reduced. Further increase in coating resulted in complete loss of oscillation in the device due to excess mass load leading to unsustainably high insertion loss.

Thus, as a trade-off between optimum sensitivity, reasonable noise performance and device stability, coating value between 2−3 mg/ml was taken as optimum value for both Ni and Fe.

2.3. Experimental setup

The experimental setup is shown in Fig. 7. It consists of the developed SAW magnetic sensor, Helmholtz coil, power supply and PC for data acquisition and display. A portable spectrum analyzer (Anritsu make) was also used to monitor the SAW frequency. The Helmholtz coil was calibrated using a standard gauss meter to provide a uniform field of upto 120 m T at a maximum current of 5 A. The SAW sensor was firmly placed on a platform between the two coil segments of Helmholtz coil. Arrangement was made to align the device in all the three directions. An interfacing software developed in-house was used for display and storage of sensor response in real time.

3. Result and discussion

Fig. 8(a) shows the response of the Fe coated SAW magnetic field sensor obtained by varying current in the magnetic coil. The sensor was alternately placed in positions so that field was parallel to x, y and z axis. The sensor frequency response showed a monotonous decrease with increase in magnetic field. The maximum sensitivity of around 6 ppm/mT was observed when the field was in x-direction. The maximum sensitivity obtained in y-direction was ∼3 ppm/mT and z-direction produced the maximum sensitivity of ∼1.2 ppm/mT. The sensor response expectedly showed a monotonous increase when the field strength was decreased. Similar results were observed when Ni coated device was tested. The results are shown in Fig. 8(b). It is obvious that the sensitivity of Ni is higher than Fe by a decent margin due to higher magnetostrictive coefficient. The results, comparing two materials in all the three axes, are tabulated in Table 3.

Even though the sensitivity in y and z directions is less compared to x direction yet it is not drastically low as is usually observed in SAW sensors based on sputtered coatings or engraved Ni IDT.

The response curves also show hysteresis phenomenon which is more pronounced in x direction for both materials because magnetic domain alignment is strong along this direction. The hysteresis was also observed more in Ni compared to Fe possibly due to higher magnetostrictive deformation.

An inherent limitation of SAW magnetostrictive thin film coated sensors is the mechanical fatigue of the film leading to the deterioration of sensor performance and shortening of life. Weak adhesion of thin film towards the surface and cyclic mechanical deforma-

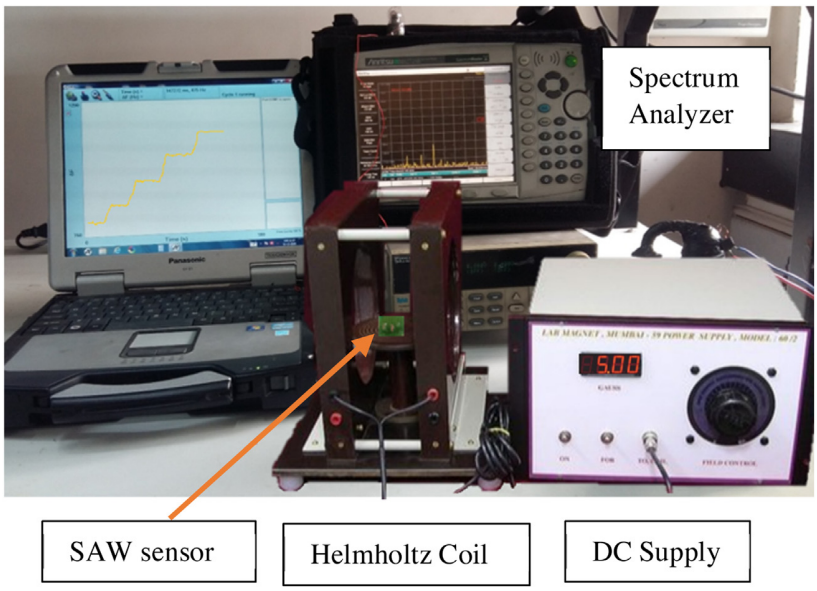
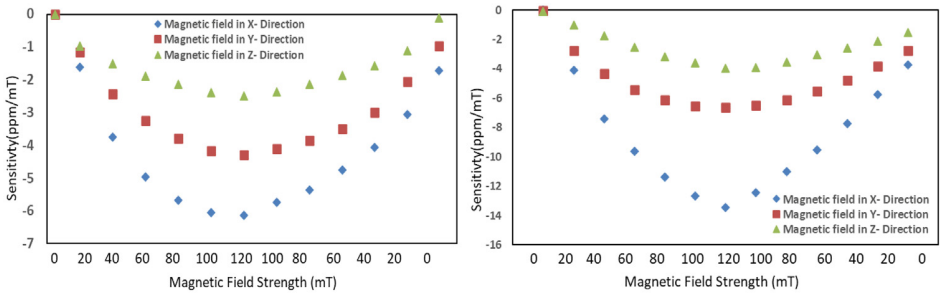


Fig. 7. Experimental setup for characterising the developed SAW magnetic sensors.



(a). Response of Fe coated SAW Sensor (b). Response of Ni coated SAW sensor

Fig. 8. (a). Response of Fe coated SAW Sensor. (b). Response of Ni coated SAW sensor.

Table 1
Lists some of the previously reported studies on SAW magnetic field sensor based on magnetostriction effect.

Device	Coating	Sensitivity	Ref.
Love wave device Quartz (ST-X+ 90°)	CoFeB (100 nm)/ SiO_2(250 nm)/ZnO(300 nm)	−420 ppm/mT	*Yang et al* [26]
ST-cut 90° X quartz, SAW 1-port resonator	Ni electrodes	2.7 ppm/mT	*Kadota et al.* [27]
Delay line on $LiNbO_3$ substrate, 42 MHz	$TbFe_2$ film	−0.27 % velocity change	*Yamaguchi et al* [28]
300 MHz Delay line SiO_2/128° YX $LiNbO_3$	FeCo coating	8.3 kHz/mT	*Wang et al* [20]
Delay line on 128°Y-X $LiNbO_3$	Ni/ZnO/IDT	3.6 ppm/mT	*Elhosni et al* [29]
433MHz 1-port resonator on ST-Quartz	PVA + Ni/Fe nanopowder	5.83 kHz/mT (13.43 ppm/mT)	*Work reported in this manuscript*

Table 2
Characterization parameters of coated SAW devices at 1–5 mg per ml of PVA solvent.

Conc. (mgs per ml of Solvent)	**Un-coated**	**1**	**2**	**3**	**4**	**5**
Freq. (MHz)	433.92	433.08	431.19	430.06	428.33	422.12
Insertion loss (dB)	−2.1	−2.96	−3.63	−4.37	−6.12	−7.91
Sensitivity (ppm/mT)	–	~1	~4	~6	~8	~9
Q-Factor	895	876	793	712	543	364
Average Noise (Hz)	1–2	10–15	25–30	30–40	100–130	>300

tions are the main reasons for film degradation and rupture [24]. Methods to improve adhesion and reduce fatigue have resulted in decrease in sensitivity. On the other hand, our method of nano sized ferromagnetic particles embedded in PVA are more or less immune to any major damage to the film due to micro deformations taking place in nano particles. In order to check the fatigue and verify the reliability of our sensors, experiments were repeated multiple times over a period of three months and performance of the sensors were found reasonably repeatable and stable. Table 4 tabulates mean frequency shift and mean deviation obtained by repeating experiments multiple times over a period of three months.

Table 3
Experimental results of developed sensors showing sensitivity in all three directions.

Description	**Device 1**	**Device 2**
Coating material	Fe	Ni
Binding material	PVA	PVA
Peak Sensitivity in increasing magnetic field strength (ppm/mT)	x-direction ~6 y-direction ~3 z-direction ~1.2	x-direction ~13.5 y-direction ~ 4 z-direction ~ 2
Peak Sensitivity in decreasing magnetic field strength (ppm/mT)	x-direction ~5 y-direction ~2.5 z-direction ~1.1	x-direction ~11.5 y-direction ~ 3.5 z-direction ~ 1.8

Table 4
Showing mean sensor response conducted for three months and mean deviation.

Coating	Field strength (mT) in x direction	No. of times experiment repeated	Mean frequency shift (kHz)	Mean Deviation $\frac{\sum\|x-\mu\|}{n}$
Ni	40	50	265.14	3.57
	80	50	546.23	4.11
	120	50	678.05	3.05
Fe	40	50	182.22	4.81
	80	50	231.41	4.20
	120	50	289.34	4.49

One of the limitations of this method is the hygroscopic nature of PVA which makes it highly sensitive to humidity and hence requires the sensor to be hermetically sealed in dry environment. High sensitivity of SAW to changes in pressure, temperature, vibrations etc. are common causes of interference. Some of these issues have been tackled by having a differential SAW arrangement with a reference sensor to cancel out the common mode effects. For better temperature compensation, the reference device also needs to be coated with equal PVA as in sensing device minus nanopowder.

Enhancement in sensitivity is another area where further insight is required. The achieved peak sensitivity of 13.5 ppm/mT using Ni nanopowder is at par or significantly better than the reported work with similar configurations (Table 1). Further enhancement in sensitivity is achievable by enhancing the sensitivity of SAW device. SAW sensitivity varies in proportion to its resonant frequency, thus a higher frequency device will have better sensitivity. However, higher frequency devices are smaller in size and will have lesser surface area for coating, thereby a reduced magnetostrictive effect. A thorough study will have to be conducted to understand the relationship. Use of Love-Wave mode-based SAW devices, having very high sensitivity (Wang [26]), can also be explored for implementing our scheme in future. At present, we need to strike a balance between frequency and size of device for optimum sensitivity and noise performance.

4. Conclusion

A novel SAW magnetic field sensor was fabricated and characterized using Ni and Fe nanopowder embedded in PVA thus forming nanocomposite film. The magnetostrictive coating on SAW using PVA passivation and PVA embedded ferromagnetic metal nanopowder has been developed for the first time to the best of our knowledge. The dry drop method used provided a simple and cost effective alternative to complex and costly coating techniques. Characterization techniques like AFM and XRD showed good quality coating could be achieved. The developed sensors showed good sensitivity and repeatability. The dependency on direction of field and concentration of metal particles in film was also seen. No effect of fatigue was observed as nano sized metal particles produced micro strains in presence of magnetic field and similar response was observed over a period of time.

SAW devices are capable of working in wireless passive mode. The feature makes these devices quite useful for measurement of current/field in remote and inaccessible locations. We are now attempting to develop a field deployable low cost wireless SAW current sensor using the technique developed through this study.

Author statement

Fahim: Conceptualization, methodology, investigation and draft preparation.
Mainuddin: Supervision, reviewing and editing.
Pooja Rajput: Characterization, Data curation.
Jitender Kumar: Experimental setup, device coating.
AT Nimal: Supervision, software.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgments

Authors are extremely thankful to Director SSPL Dr. Seema Vinayak, for permitting and supporting this work. They are also thankful to Dr. M.U. Sharma, Pallav Kumar Tripathi, Upendra Mittal and Devendra Barlewar for their support and patronage. Authors are also thankful to Ministry of Electronics and Information Technology (MietY) for their support under Visvesvaraya Scheme.

References

[1] J. Lenz, A.S. Edelstein, Magnetic sensors and their applications, IEEE Sens. J. 6 (June (3)) (2006) 631–649, http://dx.doi.org/10.1109/JSEN.2006.874493.
[2] A.C. Sreekantan, B. George, Magnetic sensors and industrial sensing applications, in: Smart Sensors and MEMS: Intelligent Sensing Devices and Microsystems for Industrial Applications, second edition, Elsevier Inc., 2018, pp. 131–150.
[3] D. Murzin, et al., Ultrasensitive magnetic field sensors for biomedical applications, Sensors (Switzerland) 20 (6) (2020), http://dx.doi.org/10.3390/s20061569.
[4] A. Roy, P. Sampathkumar, P.S. Anil Kumar, Development of a very high sensitivity magnetic field sensor based on planar Hall effect, Meas. J. Int. Meas. Confed. 156 (2020) 107590, http://dx.doi.org/10.1016/j.measurement.2020.107590.
[5] V.S. Luong, et al., Magnetoresistive performances in exchange-biased spin valves and their roles in low-field magnetic sensing applications, J. Sci. Adv.

Mater. Devices 3 (4) (2018) 399–405, http://dx.doi.org/10.1016/j.jsamd.2018.09.004.
[6] C. Sutthinet, T. Phetchakul, W. Luanatikomkul, A. Poyai, Multi-sensor diode for magnetic field and photo detection, 2015 IEEE 10th Int. Conf. Nano/Micro Eng. Mol. Syst. NEMS 2015 (2015) 179–182, http://dx.doi.org/10.1109/NEMS.2015.7147405.
[7] X. Yang, X. Zhao, Y. Bai, M. Lv, D. Wen, Two-dimensional magnetic field sensor based on silicon magnetic sensitive transistors with differential structure, Micromachines (2017), http://dx.doi.org/10.3390/mi8040095.
[8] A.L. Herrera-May, M. Lara-Castro, F. López-Huerta, P. Gkotsis, J.P. Raskin, E. Figueras, A MEMS-based magnetic field sensor with simple resonant structure and linear electrical response, Microelectron. Eng. 142 (2015) 12–21, http://dx.doi.org/10.1016/j.mee.2015.06.009.
[9] G. Singhal, Mainuddin, R. Rajesh, M.T. Beg, R.K. Tyagi, A.L. Dawar, Overview of optical techniques for characterization of high-power infrared gas lasers, IEEE Sens. J. 15 (8) (2015) 4165–4173, http://dx.doi.org/10.1109/JSEN.2015.2409114.
[10] H. Hauser, R. Steindl, J. Nicolics, A. Pohl, Miniaturized Magnetic Field Sensors Utilizing Giant Magneto-Impedance [GMI] Effect and Surface Acoustic Wave [SAW] Technology, 2001.
[11] I. Ennen, D. Kappe, T. Rempel, C. Glenske, A. Hütten, Giant magnetoresistance: basic concepts, microstructure, magnetic interactions and applications, mdpi.com (2016), http://dx.doi.org/10.3390/s16060904.
[12] L. Hao, et al., Magnetic nanoparticle detection using nano-SQUID sensors, Artic. J. Phys. D Appl. Phys. 43 (December (47)) (2010), http://dx.doi.org/10.1088/0022-3727/43/47/474004.
[13] X. Yang, Y. Li, W. Guo, W. Zheng, C. Xie, H. Yu, A new compact fluxgate current sensor for AC and DC application, IEEE Trans. Magn. 50 (11) (2014), http://dx.doi.org/10.1109/TMAG.2014.2330373.
[14] P. Ripka, K. Závěta, Chapter Three magnetic sensors: principles and applications Handbook of Magnetic Materials, 18, 2009, pp. 347–420, http://dx.doi.org/10.1016/S1567-2719(09)01803-4, no. C, Jan. 01.
[15] B. Li, O. Yassine, J. Kosel, A surface acoustic wave passive and wireless sensor for magnetic fields, temperature, and humidity, IEEE Sens. J. 15 (January (1)) (2015) 453–462, http://dx.doi.org/10.1109/JSEN.2014.2335058.
[16] F. Durani, Mainuddin, U. Mittal, J. Kumar, A.T. Nimal, Use of surface acoustic wave (SAW) for thermal conductivity sensing of gases–a review, IETE Tech. Rev. (Institution Electron. Telecommun. Eng. India) (2020), http://dx.doi.org/10.1080/02564602.2020.1819888.
[17] S. Kim, M.R. Adib, K. Lee, Development of chipless and wireless underground temperature sensor system based on magnetic antennas and SAW sensor, Sensors Actuators, A Phys. 297 (2019) 111549, http://dx.doi.org/10.1016/j.sna.2019.111549.
[18] V. Polewczyk, et al., Unipolar and bipolar high-magnetic-field sensors based on surface acoustic wave resonators, Phys. Rev. Appl. 8 (August (2)) (2017), http://dx.doi.org/10.1103/PhysRevApplied.8.024001.
[19] Fahim, Mainuddin, U. Mittal, J. Kumar, A.T. Nimal, Novel SAW CWA detector using temperature programmed desorption, IEEE Sens. J. (1) (2020), http://dx.doi.org/10.1109/JSEN.2020.3042766.
[20] W. Wang, Y. Jia, X. Liu, S. He, in: Development of a Novel SAW Current Sensor Based on the Magnetostrictive Effect, 2015, http://dx.doi.org/10.1109/ULTSYM.2015.0354, Nov.
[21] M. Kadota, S. Ito, Y. Ito, T. Hada, K. Okaguchi, SAW magnetic sensors composed of various Ni electrode structures on quartz, IEEE International Ultrasonics Symposium, IUS (2011) 805–809, http://dx.doi.org/10.1109/ULTSYM.2011.0197.
[22] B. Li, H. Al, J. Kosel, Surface acoustic wave based magnetic sensors, in: Modeling and Measurement Methods for Acoustic Waves and for Acoustic Microdevices, InTech, 2013.
[23] M. Eslamian, F. Soltani-Kordshuli, Development of multiple-droplet drop-casting method for the fabrication of coatings and thin solid films, J. Coatings Technol. Res. 15 (March (2)) (2018) 271–280, http://dx.doi.org/10.1007/s11998-017-9975-9.
[24] Y. Jia, et al., Fatigue characteristics of magnetostrictive thin-film coated surface acoustic wave devices for sensing magnetic field, IEEE Access 8 (2020) 38347–38354, http://dx.doi.org/10.1109/ACCESS.2020.2976052.
[25] E. Hristoforou, A. Ktena, Magnetostriction and magnetostrictive materials for sensing applications, J. Magn. Magn. Mater. 316 (2) (2007) 372–378, http://dx.doi.org/10.1016/j.jmmm.2007.03.025, SPEC. ISS.
[26] Y. Yang, et al., Enhanced performance love wave magnetic field sensors with temperature compensation, IEEE Sens. J. 20 (19) (2020) 11292–11301, http://dx.doi.org/10.1109/JSEN.2020.2998826.
[27] M. Kadota, S. Ito, Sensitivity of surface acoustic wave magnetic sensors composed of various Ni electrode structures, J. Appl. Phys. 51 (7) (2012), http://dx.doi.org/10.1143/JJAP.51.07GC21, PART2.
[28] M. Yamaguchi, K. Hashimoto, H. Kogo, M. Naoe, Variable SAW delay line using amorphous TbFe&;lt;inf&;gt;2&;lt;/inf&;gt;film, IEEE Trans. Magn. 16 (5) (1980) 916–918 [Online]. Available: http://ieeexplore.ieee.org/document/1060738/.
[29] M. Elhosni, et al., Magnetic field SAW sensors based on magnetostrictive-piezoelectric layered structures: FEM modeling and experimental validation, Sensors Actuators, A Phys. 240 (April) (2016) 41–49, http://dx.doi.org/10.1016/j.sna.2015.10.031.
[30] A.T. Nimal, M. Singh, U. Mittal, R.D.S. Yadava, A comparative analysis of one-port Colpitt and two-port Pierce SAW oscillators for DMMP vapor sensing, Sensors Actuators, B Chem. 114 (1) (2006) 316–325, http://dx.doi.org/10.1016/j.snb.2005.05.021.
[31] Fahim, Mainuddin, U. Mittal, J. Kumar, A.T. Nimal, M.U. Sharma, Single chip readout electronics for SAW based gas sensor systems, Proc. IEEE Sens. 2017 (December) (2017) 1–3, http://dx.doi.org/10.1109/ICSENS.2017.8233886.
[32] A. Hausoel, et al., Local magnetic moments in iron and nickel at ambient and Earth's core conditions, Nat. Commun. 8 (July) (2017), http://dx.doi.org/10.1038/ncomms16062.
[33] M.K. Shobana, V. Rajendran, K. Jeyasubramanian, N. Suresh Kumar, Preparation and characterisation of NiCo ferrite nanoparticles, Mater. Lett. 61 (13) (2007) 2616–2619, http://dx.doi.org/10.1016/j.matlet.2006.10.008.
[34] S. Kayal, R.V. Ramanujan, Doxorubicin loaded PVA coated iron oxide nanoparticles for targeted drug delivery, Mater. Sci. Eng. C 30 (3) (2010) 484–490, http://dx.doi.org/10.1016/j.msec.2010.01.006.
[35] H. Farahani, R. Wagiran, M.N. Hamidon, Humidity sensors principle, mechanism, and fabrication technologies: a comprehensive review, Sensors Basel (Basel) 14 (5) (2014).
[36] M. Aslam, M.A. Kalyar, Z.A. Raza, Polyvinyl alcohol: a review of research status and use of polyvinyl alcohol based nanocomposites, Polym. Eng. Sci. 58 (December (12)) (2018) 2119–2132, http://dx.doi.org/10.1002/pen.24855, John Wiley and Sons Inc.
[37] P.K. Khanna, N. Singh, S. Charan, V.V.V.S. Subbarao, R. Gokhale, U.P. Mulik, Synthesis and characterization of Ag/PVA nanocomposite by chemical reduction method, Mater. Chem. Phys. 93 (1) (2005) 117–121, http://dx.doi.org/10.1016/j.matchemphys.2005.02.029.
[38] D.K. Rana, S.K. Singh, S.K. Kundu, R.J. Choudhary, S. Basu, Electrical and magnetic properties of polyvinyl alcohol-cobalt ferrite nanocomposite films, Bull. Mater. Sci. 41 (4) (2018), http://dx.doi.org/10.1007/s12034-018-1608-0.
[39] A. Kittmann, et al., Sensitivity and noise analysis of SAW magnetic field sensors with varied magnetostrictive layer thicknesses, Sensors Actuators, A Phys. 311 (2020), http://dx.doi.org/10.1016/j.sna.2020.111998.

Biographies



Fahim received BE degree from National Institute of Technology (NIT), Srinagar in 2000 and M.Tech in Instrument Technology from Indian Institute of Technology (IIT) Delhi in 2012. He is working as scientist in Solid State Physics Laboratory, Delhi. He is currently pursuing PhD from Jamia Millia Islamia University, Delhi. His research interests include Surface Acoustic Wave (SAW) sensor based chemical Detectors, FPGA design, sensor interface electronics and Instrumentation.



Mainuddin received Bachelor's degree in Electronics and Communication Engineering from Jamia Millia Islamia. He received M.E. degree from the Department of Electronics and Communication Engineering, Delhi College of Engineering in 2003. He completed his Ph.D. in 2008 from Jamia Millia Islamia, New Delhi. He worked as a Senior Research Scientist at Laser Science and Technology Centre, DRDO, New Delhi for more than 16 years. He is presently working as a Professor at the department of Electronics and Communication Engineering, Jamia Millia Islamia, New Delhi. He has over 100 research publication in journals/ conferences. His research interests include Optoelectronics, Optical diagnostics, Sensors and Instrumentation and high power lasers.



Pooja Rajput received her B.Tech (2013) in electronics and communication from Uttar Pradesh Technical University, Lucknow (U.P.) and M.Tech (2015) in VLSI Design from Banasthali University, Rajasthan, India. She is working as a research fellow at Solid State Physics Laboratory (DRDO), Delhi since 2016 and pursuing her Ph.D. from Centre for Nanoscience and Nanotechnology, Jamia Millia Islamia, Delhi. Her research interests are in temperature study of SAW device, development of thin films for sensor coatings and temperature stability analysis of SAW devices with various coating materials for sensor applications.



Jitender Kumar received his BSc (2001) in Life science from University of Meerut, Meerut, and MSc (2004) in Organic Chemistry from University of Meerut, Meerut. He is working as a scientist at Solid State Physics Laboratory, Delhi since 2008. He is involved in the development & characterization of Gas Chromatography and SAW based chemical sensor. He is also involved in the development of vapour sampler and Preconcentrator setup for toxic industrial chemicals.



A.T. Nimal received his BSc (1991) and MSc (1993) in Physics from Government Arts College, Bharathiar University, Coimbatore and PhD (2004) in Physics from National Physical Laboratory, Delhi, and University of Delhi, Delhi. He is working as a scientist at Solid State Physics Laboratory, Delhi, since 1997. Presently he is involved in the development of SAW based Electronic Nose for CW agents. He was also involved in the design and development of SAW devices such as delay lines, dispersive delay lines, resonators and filters for the above sensors and signal processing. He has also worked in the field of high TC superconducting thin films and transition edge bolometer IR sensors during his PhD.

ARTICLE IN PRESS

Materials Today: Proceedings xxx (xxxx) xxx

Contents lists available at ScienceDirect

Materials Today: Proceedings

journal homepage: www.elsevier.com/locate/matpr

A Study of thermoelectric properties of Hf-doped RhTiP Half-Heusler alloy

Shivam Tyagi, Shivani Arora, Mukhtiyar Singh *

Department of Applied Physics, Delhi Technological University, Delhi 110042, India

ARTICLE INFO

Article history:
Received 19 April 2021
Received in revised form 20 May 2021
Accepted 26 May 2021
Available online xxxx

Keywords:
Heusler thermoelectric
DFT
Boltztrap
RhTiP
GGA

ABSTRACT

In this work, we present a *first-principles* based investigation to analyse the effect of Hafnium (Hf) doping on thermoelectric properties of half-Heusler alloy RhTiP. In recent times, Hf based Heuser alloys are proved to be good thermoelectric (TE) materials. So, its interesting to see the effect of Hf, as dopant, on the newly propsed RhTiP. The energy bandgap of pure RhTiP is 0.810 eV. As we start doping, the bandgap first decrease for $RhTi_{0.75}Hf_{0.25}P$ and then increase liearly for higher doping concentrations. We find that the material shows the high thermoelectric performance at 100% doping of Hf in place of Ti (RhHfP) at room temperature. The maximum value of power factor and electrical conductivity at room temperature are estimated, respectively, as $0.19688x10^{12} WK^{-2}m^{-1}s^{-1}$ and $3.2058x10^{18}\Omega^{-1}m^{-1}s^{-1}$ which belong to RhHfP . The maximum *ZT* value at room temperature is found to be 0.7950 at room and it is too for RhHfP. We estimate that RhHfP can become a good thermoelectric material amongst all the studied concentrations for room temperature applications.
© 2021 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the 1st International Conference on Computations in Materials and Applied Engineering – 2021.

1. Introduction

The exponential upsurge in demand for energy and continuous degradation of natural energy resources like depletion of fossil fuel are the significant areas of concern because a high portion of usable energy comes from burning the fossil fuels. These resources are present in the limited quantity on earth and this resulted in an extensive research worldwide on environment-friendly and efficient ways to find the non-renewable sources of energy. One of such ways is efficient thermoelectric materials (TE) that can convert waste heat into useful electricity. These TE materials, in addition to their use into power generation devices, can also be implemented into refrigeration devices for cooling applications as well [1–3]. The thermoelectric performance of a material can be determined using a parameter known as the figure of merit denoted by *ZT* and given as

$$ZT = \frac{S^2\sigma T}{\kappa} \tag{1}$$

where *S* is the Seeback coefficient, σ is the electrical conductivity, *T* is the absolute temperature, and κ is known as the thermal conductivity given as

$$\kappa = \kappa_e + \kappa_l, \tag{2}$$

where $\kappa_e$is the electronic thermal conductivity and κ_l is the lattice thermal conductivity. A material having a high figure of merit (*ZT*) has good thermoelectric properties. The dependence of *ZT* on various parameters can be easily determined from equation (1). The electronic thermal conductivity given by the Wiedemann-Franz Law [4], that is, $\kappa_e = L\sigma T$,where, L is the Lorentz constant ($\approx 2.4 * 10^{-8} J^2K^2C^{-2}$).

Half-Heusler alloys have gained considerable interest in the present times due to their high-temperature stability, comparatively high Seebeck coefficient, and relatively high *ZT* value [5–11]. Majority of these alloys have a bandgap between 0.1 and 1.0 eV [12–13], which is also a merit of these systems over conventional semiconductore. It can be seen from equation (1) that the ZT value of a material can be enhanced either by decreasing the value of lattice thermal conductivity or by increasing electrical conductivity but these two perperties are correlated with each other and its is very difficult to manipulate them in a favorable way. However, there are other various ways have been proposed to

* Corresponding author.
E-mail address: mukhtiyarsingh@dtu.ac.in (M. Singh).

https://doi.org/10.1016/j.matpr.2021.05.570
2214-7853/© 2021 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the 1st International Conference on Computations in Materials and Applied Engineering – 2021.

Please cite this article as: S. Tyagi, S. Arora and M. Singh, A Study of thermoelectric properties of Hf-doped RhTiP Half-Heusler alloy, Materials Today: Proceedings, https://doi.org/10.1016/j.matpr.2021.05.570

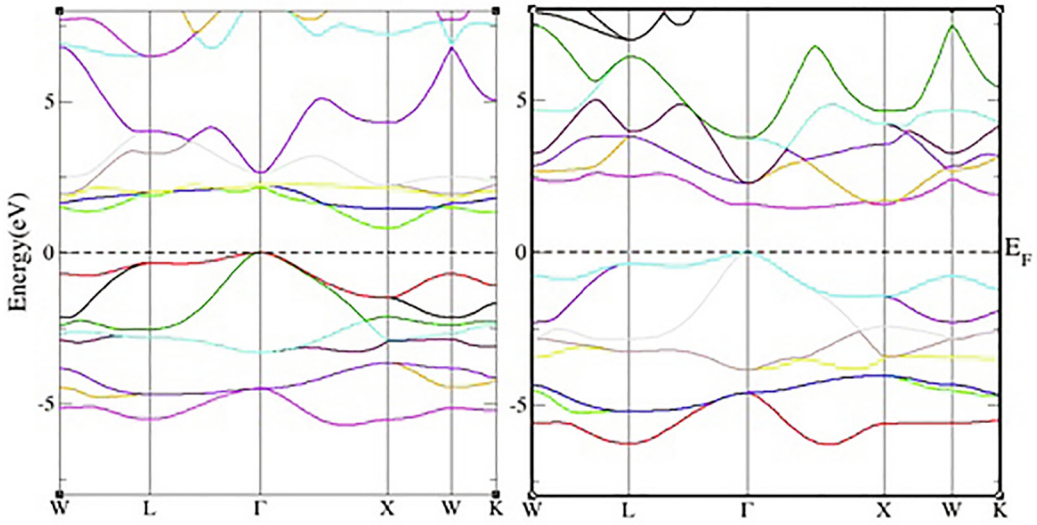


Fig 1. (a) Bandstructure of RhTiP; and of (b) RhHfP.

enhance the TE performance of these materials. The most efficient such ways include tuning of band structure by applying the strain or pressure and by doping of suitable elements [14–19].

In the family of ABX type Half heusler alloy, RhTiP is a newly found material that is proven to be stable in previous study [20], and it is yet to be developed experimentally. It can be used as an excellent thermoelectric material since it has a pretty high-power factor and Seebeck coefficient. It is a ductile material with a bonding nature to be weakly ionic [20], so the property of the material to withhold external pressure and strain is relatively low in this case. Due to the presence of weak ionic bond, its disassociation energy should also be low; hence it is not on the problematic side to replace a specific atom from the stoichiometric composition. Moreover, there are already many good thermoelectric materials with the same configuration TaCoSn [21], ScPtSb [22]. This motivates us to work on the possibility of enhancing the *ZT* value of RhTiP half-heusler alloy, which eventually increases its thermoelectric performance. So, we in the presented work, we have performed the electronic and thermoelectric calculation of Hf doped RhTiP. The main reason for using Hf as a doping agent is its chemical stability and pretty high boiling and melting point so that it can withstand high temperatures. The aim of this study is to get high stable material with good thermoelectric properties at different concentrations of Hf. To best of our knowledge this work has not been yet undertaken and we believed that this work stimulus forther experimental research on this system.

2. Computational details

The structural, electronic, and thermal transport properties of pristine RhTiP and Hf-doped RhTiP were investigated using Density Functional Theory (DFT) [23–24] based Full Potential Linear Augmented Plane Wave (FPLAPW) method as implemented in Wein2K code [25]. Generalized gradient approximation (GGA) formalism within the Perdew-Burke-Ernzerhof (PBE) parametrization was used to handle the exchange–correlation potentials [26]. Plane wave cut-off parameters were decided by $R_{MT}K_{max}$ = 7. A uniform K-mesh of 20 × 20 × 20 was used for 8000 k points Brillouin zone integration which was done by using the original Monkhorst-Pack's scheme in the irreducible Brillouin zone (IBZ). The energy convergence criterion was fixed to 0.00001, and charge convergence was monitored accordingly. The transport properties were obtained using the BoltzTrap code [27], which solved the semiclassical Boltzmann transport equation while having a constant value of relaxation time τ. This code was used to determine factors like power factor, Seebeck coefficient, electrical conductivity, etc. The power factor $S^2\sigma$ is given in units of τ, where τ is relaxation time [28]. We were constructed a 1x1x1 supercell of 12 atoms in order to simulate the various doping concentration.

3. Results and discussion

First of all, we have performed the structural optimization of RhTiP, and it is found to be stable in the cubic semiconducting phase with a lattice constant of 5.76 Å. This value of lattice constant is in good agreement with what is available in the literature. The positions of the atoms inside crystal are taken to be (0.25, 0.25, 0.25) for Rh atom, (0.5, 0.5, 0.5) for Ti atom and (0, 0, 0) for P atom. The rest of the calculations have been performed on the optimized lattice constant.

The semiconducting nature of all the concentrations in $RhTi_{1-x}Hf_xP$ (x = 0.25, 0.50, 0.75, 1) have found to be intact. We have obtained the energy bandgap value for RhTiP as 0.81 eV, which is in fair agreement with the previous study, 0.83 eV [20]. The energy dispersion has been plotted, Fig. 1, along with high symmetry, points $W \rightarrow L \rightarrow \Gamma \rightarrow X \rightarrow W \rightarrow K$.

We have observed that pure RhTiP has an indirect bandgap whose valence maxima were at Γ and the conduction minima were at X high-symmetry point, while during the process of doping, the nature of the bandgap was tuned. The values of the bandgap for all studied concentrations are listed in Table 1.

As shown in Fig. 2, the density of states (DOS) for $RhTi_{0.75}Hf_{0.25}P$ attains its maximum value in valence band around −3 eV and

Table 1
Resultant alloys after doping and their respective bandgap values

Alloys	Bandgap (in eV)
RhTip	0.815 (0.830) *
RhTiP	0.810
$RhTi_{0.75}Hf_{0.25}P$	0.797
$RhTi_{0.5}Hf_{0.5}P$	0.877
$RhTi_{0.25}Hf_{0.75}P$	1.078
RhHfP	1.464

* Other's work [20].

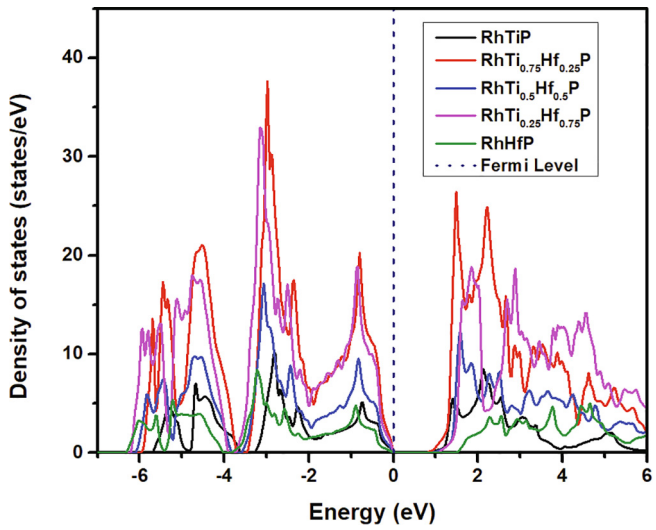


Fig 2. Density of states plot of $RhTi_{1-x}Hf_xP$ (x = 0.25, 0.50, 0.75, 1).

in c onduction band peak occur at approx 1.5 eV. Here both 25% and 75% doping concentrations possess the highest DOS value in both valence and conduction band, while both RhTIP and RhHfP have the minimum value amongst them. The young's (E), shear (G), and bulk modulus (B) of this material were found in earlier studies [20] and are 177.72, 66.21, 188.58, respectively, which shows the stability of the materials under investigation.

Let us now move to understand the transport properties of $RhTi_{1-x}Hf_xP$ (x = 0.25, 0.50, 0.75, 1). As a starting point, we first discussed the thermal-transport properties of pure RhTiP half-Heusler alloy. The value of the Seebeck coefficient of RhTiP has been found to be 247.1 $\frac{\mu V}{K}$ at room temperature, and this value approached up to 247.7 $\frac{\mu V}{K}$ at 1200 K (as depicted in Fig. 3(a)).

The maximum power factor of the parent system has been found out to be $1.44 \times 10^{12} WK^{-2}m^{-1}s^{-1}$ at 1200 K temperature and at room temperature, its value has been $1.823 \times 10^{11} WK^{-2}m^{-1}s^{-1}$. Finally, the ZT value of pure RhTiP has found to be maximum at room temperature, that is, 0.77479. After that, we have doped this pure RhTiP with Hf, which leads to a chemically stable compounds and it has been established that the alloys doped with an impurity such as Hf and Zr are found to be chemically stable [28–29]. Doping of Hf has been done such as we got four different concentrations $RhTi_{0.75}Hf_{0.25}P$, $RhTi_{0.5}Hf_{0.5}P$, $RhTi_{0.25}Hf_{0.75}P$, and finally RhHfP (with complete replacement of Ti with Hf).

The value of the Seebeck coefficient has observed to decrease with an increase in doping percentage and, as shown in Fig. 3(a),

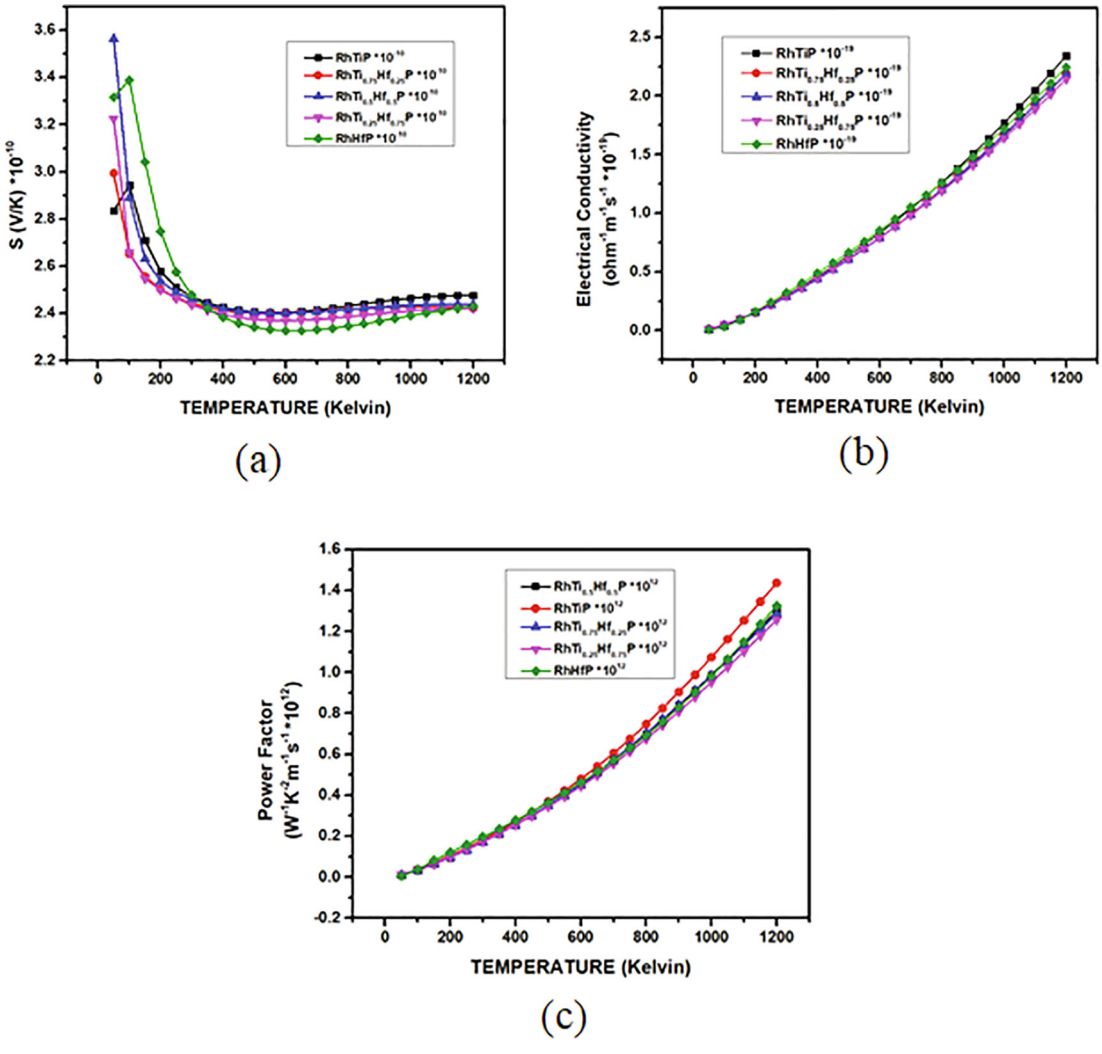


Fig 3. (a) Plot of Seebeck coefficient for different doping concentrations with respect to temperature, (b) Plot of electrical conductivity for different doping concentrations with respect to temperature, (c) Plot of power factor of different doping concentrations versus temperature.

we can clearly see that it first decreased in all cases and then increased with respect to temperature. The maximum value of the Seebeck coefficient is 356.407 $\mu V/K$ of $RhTi_{0.5}Hf_{0.5}P$ at 50 K, and the minimum for alloy $RhTi_{0.25}Hf_{0.75}P$ at 650 K (232.712$\mu V/K$). Interestingly the value of electrical conductivity increased with the increase in temperature (as shown in Fig. 3 (b)) but here also the maximum value of electrical conductivity belongs to pure RhTiP at 1200 K ($2.34x10^{19}\Omega^{-1}m^{-1}s^{-1}$)..

As in Fig. 3(c), we can see the power factor also increased with increasing in temperature (as the formula suggests PF = $S^2\sigma$) just like the electrical conductivity (also electrical conductivity of pure alloy is maximum at almost all temperature range that might be because of the fact that radius of Ti is smaller than that of doping element Hf so charge carriers have a lesser space to move around and that may result in decreasing its electrical conductivity), the maximum value of power factor belongs to pure alloy RhTiP (Fig. 3(c)) that is, approximately around $1.44 \times 10^{12}WK^{-2}m^{-1}s^{-1}$ at 1200 K and in the doped systems, the maximum value of the power factor is of RhHfP, which is $1.321x\ 10^{12}WK^{-2}m^{-1}s^{-1}$. The calculated maximum values of power factor for other doped concentrations such as $RhTi_{0.5}Hf_{0.5}P$, $RhTi_{0.75}Hf_{0.25}P$, and $RhTi_{0.25}Hf_{0.75}P$ have found to be $1.300 \times 10^{12}WK^{-2}m^{-1}s^{-1}$, $1.285 \times 10^{12}WK^{-2}m^{-1}s^{-1}$and $1.25704 \times 10^{12}WK^{-2}m^{-1}s^{-1}$, respectively. At room temperature, the power factor for RhHfP has been found to be maximum ($0.19688x10^{12}WK^{-2}m^{-1}s^{-1}$) while on the other hand, at room temperature $RhTi_{0.75}Hf_{0.25}P$ have a minimum power factor ($0.17152\ x10^{12}WK^{-2}m^{-1}s^{-1}$) which is still not that much low but on the higher side, so even at the minimum value, we can say that this material can work as a good thermoelectric material.

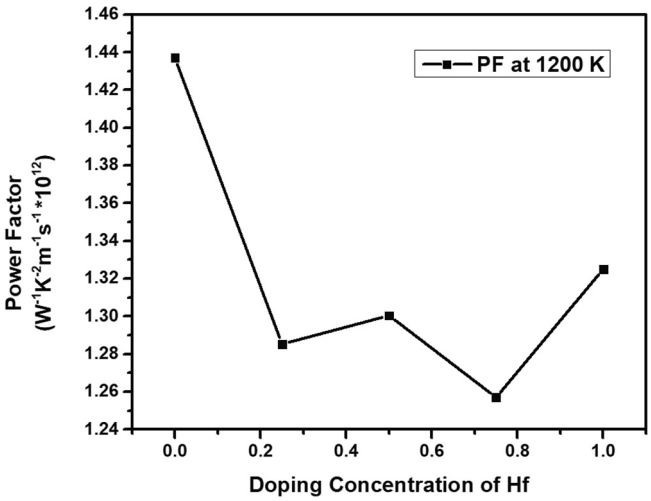


Fig 4. Plot of the maximum value of PF versus doping concentration.

The variation of the power factor with doping percentage of Hf is shown in Fig. 4.

The figure of merit (*ZT*) is maximum at 50 K for all doping concentrations (as shown in Fig. 5(a)), and amongst them, the highest *ZT* value is 0.87327, which is of $RhTi_{0.5}Hf_{0.5}P$. The ZT values for RhTiP, $RhTi_{0.75}Hf_{0.25}P$, $RhTi_{0.25}Hf_{0.75}P$ and RhHfP have been found out to be 0.75406, 0.84247, 0.86763, 0.80341, respectively. We have also found that the value of maximum value of *ZT* to be 0.87327 at 50 K of 50% doping concentration, but 50 K is a very low temperature which is of almost no use. So having a high value on that temperature couldn't quite help us. Hence, if we look towards a more realistic one, then at around 250 K, we observed that RhHfP has the maximum *ZT* value, which is approx. 0.81. All the variations in ZT value at different temperatures can be seen in Fig. 5(b). We have also observed the increase in *ZT* value of parent alloy attain peak at somewhere about 100 K, and then again it decreases slowly. All the other alloys also have their respective peaks at different temperatures. At room temperature, the *ZT* values are 0.77479, 0.76701, 0.77137, 0.76892, and 0.79534 for 0%, 25%, 50%, 75% and 100% doping concentrations of Hf, respectively. In this, we can easily see that at room temperature RhHfP possess the maximum ZT value amongst all other compounds. We can see that the variation of the maximum value of *ZT* is not linear with doping concentration; it decreased for 25% doping, increase for 50%, then again decrease to 75% and finally attain the highest peak at 100% doping of Hf in place of Ti at 300 K.

4. Conclusion

We have carefully calculated the thermoelectric properties of half-Heusler alloy RhTiP by doping different Hf concentrations using first-principles calculations. We have found that the bandgap of RhTiP increases with the effect of doping except for $RhTi_{0.5}Hf_{0.5}P$, where it decreased than the parent value and reached its maximum value of 1.464 eV for 100% doping of Hf in place of Ti. At room temperature, the power factor of RhHfP is maximum amongst all, which has found to be 1.9688 $*10^{11}$, and the maximum ZT value at room temperature is of again RhHfP which has the value of 0.79534. The $RhTi_{1-x}Hf_xP$ has a maximum power factor of

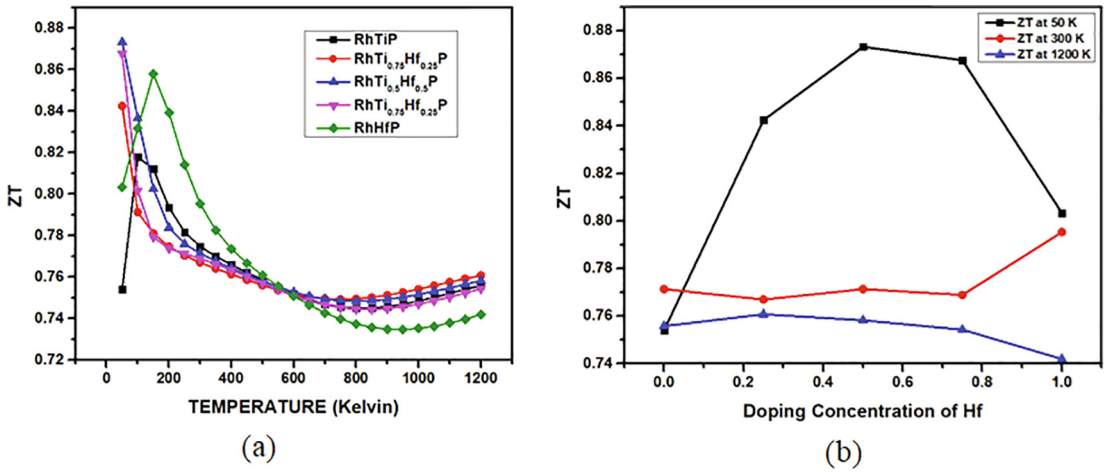


Fig 5. (a) Plot of ZT value at different temperatures, (b) Plot of ZT at 50 K, 300 K, and 1200 K versus the doping concentration.

$1.43712x10^{12} WK^{-2}m^{-1}s^{-1}$ which is relatively high and makes it a very promising alloy deserving experimental synthesis. Also, the maximum value of ZT has value 0.87327 for $RhTi_{0.5}Hf_{0.5}P$ at 50 K. Other values are not very much less than 1, which proves that all these studied systems hold a high potential to be promising thermoelectric materials. However, a detail investigation of their lattice thermal conductivity is still lacking, and it will be quite interesting to work on it.

CRediT authorship contribution statement

Shivam Tyagi: Investigation, Data curation, Writing - original draft, Software. **Shivani Arora:** Data curation, Writing - original draft, Validation. **Mukhtiyar Singh:** Supervision, Conceptualization, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

One of the authors (Mukhtiyar Singh) acknowledges the financial support from University Grant Commission (UGC) India provided through START-UP GRANT (N0. F.30-416/2018 (BSR), dated 18/09/2018), for setting the computational facilities at Department of Applied Physics, Delhi Technological University, Delhi, India.

References

[1] J.R. Sootsman, D.Y. Chung, M.G. Kanatzidis, Angew. Chem. Int. Ed. 48 (2009) 8616.
[2] S. Ouardi, G.H. Fecher, B. Balke, X. Kozina, G. Stryganyuk, C. Felser, Phys. Rev. B 82 (2010) 085108.
[3] L. Müchler, F. Casper, B. Yan, S. Chadov, C. Felser, Phys. Status Solidi RRL 7 (2013) 91.
[4] G.V. Chester, A. Thellung, Proc. Phys. Soc. 77 (1961) 1005.
[5] S.R. Clup, S.J. Poon, N. Hickman, T.M. Tritt, J. Blumm, Appl. Phys. Lett. 88 (2006) 042106.
[6] D.Y. Jung, K. Kurosaki, C.E. Kim, H. Muta, S. Yamanaka, J. Alloys Compd. 489 (2010) 328.
[7] Q. Shen, L.D. Chen, T. Goto, T. Hirai, J. Yang, G.P. Meisner, C. Uher, Appl. Phys. Lett. 79 (2001) 4165.
[8] S. Sakurada, N. Shutoh, Appl. Phys. Lett. 86 (2005) 082105.
[9] K. Ciesielski, K. Synoradzki, I. Veremchuk, P. Skokowski, D. Szymański, Y.u. Grin, D. Kaczorowski, Phys. Rev. Appl. 14 (2020) 054046.
[10] J. Yu, Y. Xing, C. Hu, Z. Huang, Q. Qiu, C. Wang, K. Xia, Z. Wang, S. Bai, X. Zhao, L. Chen, T. Zhu, Adv. Ene. Mater. 10 (2020) 2000888.
[11] Y. Hu, Y. Jin, G. Zhang, Y. Yan, RSC Adv. 10 (2020) 28501–28508.
[12] X. Yan, W.S. Liu, S. Chen, H. Wang, Q. Zhang, G. Chen, Z.F. Ren, Adv. Energy Mater. 3 (2013) 1195–1200.
[13] X. Zhang, Y. Wang, Y. Yan, C. Wang, G. Zhang, Z. Cheng, F. Ren, H. Deng, J. Zhang, Sci. Rep. 6 (2016) 33120.
[14] W. Xie, A. Weidenkaff, X. Tang, Q. Zhang, J. Poon, T.M. Tritt, Nanomaterials 2 (2012) 379.
[15] S.W. Kim, Y. Kimura, Y. Mishima, Intermetallics 15 (2007) 349.
[16] D.I. Bilc, P. Ghosez, Phys. Rev. B 83 (2011) 205204.
[17] D. Zhao, M. Zuo, Z. Wang, X. Teng, H. Geng, Func. Mater. Lett. 7 (2014) 1450032.
[18] Y. Kimura, Y. Tamura, T. Kita App, Phys. Lett. 92 (2008) 012105.
[19] R.O. Suzuki, T. Kyono, J. Alloys Compd. 377 (2004) 38.
[20] M.K. Bamgbose, P.O. Adebambo, G.T. Solola, G.A. Adebayo, J. Mater. Sci. Engn. B 264 (2021) 114987.
[21] E. Haque, M.A. Hossain, Results in Physics 10 (2018) 458–465.
[22] M.J. Winiarski, K. Bilińska, K. Ciesielski, D. Kaczorowski, J. Alloys Compd. 762 (2018) 901–905.
[23] P. Hohenberg, W. Kohn, Phys Rev. 136 (1964) B864.
[24] R.G. Parr, Density Functional Theory of Atoms and Molecules, in: K. Fukui, B. Pullman (Eds.), Horizons of Quantum Chemistry. Académie Internationale Des Sciences Moléculaires Quantiques / International Academy of Quantum Molecular Science, vol 3, Springer, Dordrecht, 1980, https://doi.org/10.1007/978-94-009-9027-2_2.
[25] P. Blaha, K. Schwarz, G.K.H. Madsen, D. Kvasnicka, J. Luitz, WIEN2k, An Augmented Plane Wave + Local Orbitals Program for Calculating Crystal Properties, Karlheinz Schwarz, Karlheinz Schwarz, Techn. Universitat Wien, Wien, Austria, 2001.
[26] K.B. Perdew, M. Ernzerhof, Phys. Rev. Lett. 77 (1996) 3865.
[27] G.K.H. Madsen, D.J. Singh, Computer Physics Communications 175 (2006) 67–71.
[28] A. ElKhouly, A. Novitskii, A.M. Adam, A. Sedegov, A. Kalugina, D. Pankratova, D. Karpenkov, V. Khovaylo, J. Alloys Compd. 820 (2020) 153413.
[29] A. Tavassoli, F. Failamani, A. Grytsiv, G. Rogl, P. Heinrich, H. Müller, E. Bauer, M. Zehetbauer, P. Rogl, Acta materialia 135 (2017) 263–276.

Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021)
IEEE Xplore Part Number: CFP21K74-ART; ISBN: 978-0-7381-1327-2

A Survey on Artificial Neural Network based Stock Price Prediction Using Various Methods

Shailender Kumar
Department of Computer Science and Engineering
Delhi Technological University
New Delhi, India
shailenderkumar@dce.ac.in

Rounak Gupta
Department of Computer Science and Engineering
Delhi Technological University
New Delhi, India
rounak176@gmail.com

Puneet Kumar
Department of Computer Science and Engineering
Delhi Technological University
New Delhi, India
pppuneetkumar12345@gmail.com

Naman Aggarwal
Department of Computer Science and Engineering
Delhi Technological University
New Delhi, India
naman00198@gmail.com

2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) | 978-1-6654-1272-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICCS51141.2021.9432329

Abstract— **Stock price prediction has been a big challenge since many years due to the unpredictable nature of the market with around 200 transactions per second, making it hard to predict the stock price. Stock market also known as share market or equity market is a cluster of sellers and buyers of stock. Big and private companies sell a part of their company to investors in order to generate crowdfunding. Since the stock market is quite sensitive to economic sentiments and rapid changes, the main objective of the stock market is to develop new techniques to forecast higher profits for stock prices. This paper aims to examine different deep learning methods by which the future stock prices can be accurately predicted.**

Keywords— Stock Market, Stock Exchange, Investors, Crowdfunding, Stock Index, SEBI, NIFTY50, BSE SENSEX.

I. INTRODUCTION

The stock market's fluctuation has an effect on an individual's life in both financial and non-financial ways. For stock investors, associate correct prediction of stock worth movement could yield profits. It involves a lot of industries including communication, finance, household appliances, medicine, automobile, civil aviation and other industries. The fluctuation in price of the stock market is closely related to the development of the national economy. Therefore, the study of the stock market is of great significance. It is quite essential to gather the necessary data to learn stock movement trends and trading habits to build an efficient market trading strategy. Many researchers have shown that social media knowledge is a valuable resource for assessing the trends and prices of investors. Investors often try to track the risks in real-time to maximize the return on investment. Forecasting allows them to reduce the risk and secure the trade between buyers and sellers. This paper discusses various methods in machine learning for prediction of stock market movement.

In 1992 SEBI Act established the Securities and Exchange Board of India (SEBI). Its principal motive is to regulate stock exchange in India. Its basic functionality includes protecting the investor's interest and promoting the regulating the Indian securities market. SEBI regulations govern the Indian securities market in which the financial intermediaries, either domestic or foreign, participate by the permission of their respective regulators.

Two major stocks in India are, NIFTY and SENSEX.

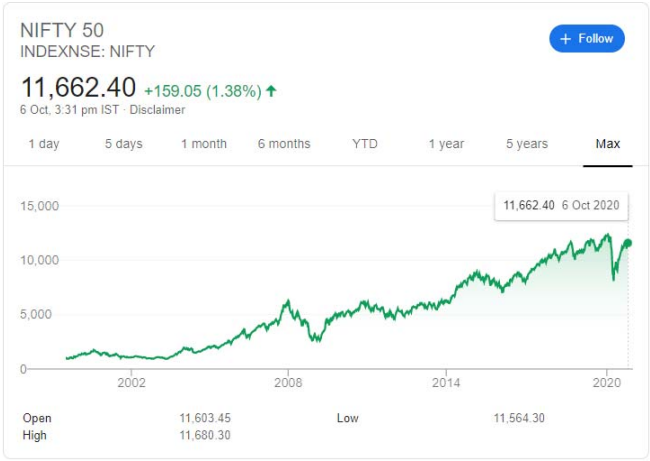


Figure 1 : Graph representing the value stock NIFTY 50 against time.



Figure 2 : Graph representing the value stock BSE SENSEX against time.

Nifty 50 and BSE Sensex are two major benchmark of Indian stock market. Nifty 50 (Fig 1), represents the average of top 50 Indian companies, whereas BSE Sensex (Bombay

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:34:22 UTC from IEEE Xplore. Restrictions apply.

Stock Exchange) (Fig 2), represents 30 well established companies listed on Bombay stock exchange.

II. Dataset used and evaluting parameters used

The stock price of '**ADANIPORT**' from December 2007 to October 2020 is the dataset used in this project. The dataset used was fetched from Quandl.

Performance of the described models is evaluated using various evaluation metrics.

A. *Root Mean Square(RMS)*

The RMS value is determined by taking the squared function mean's square root. A special case of the generalized mean with exponent 2 is the quadratic mean or RMS. RMS can also be defined as constantly changing function as to the integral of the squares of the instantaneous values in a loop.

B. *R-Squared (R2) Score*

R-squared or R2 is a mathematical variable that reflects a regression model's level of fit. For r-square, the optimal value is 1. The closest the r-square value is to 1, the more the model suits The R-square is the difference between the residual sum of squares (SSres) and the final sum of squares (SStot).

C. *Mean Absolute Error(MAE)*

The level of error in our calculations is known as absolute error. It's the disparity between what's calculated and what "reality." The absolute mistake of a scale is 90 pounds – 89 pounds = 1 pound if it reads 90 pounds but you know your true weight is 89 pounds.
The sum of all absolute errors is called the **Mean Absolute Error (MAE).**

III. Different Techniques Used

A. *Moving Average Method*

Moving Average[1] is technique in which the future value is predicted by analyzing the average of different subset out of the full data set. For example, take an imaginary scenario where we have past data of some event and want to predict the future coming event outcome using the given past data. Here we use the past event data and take the average of some definite interval and predict the future outcome on the basis of the average. In this technique we use the moving window technique where we use some interval data and predict the next upcoming event data and then we use the same predicted data and remove the oldest data we used in the last prediction and predict the next event and so on.

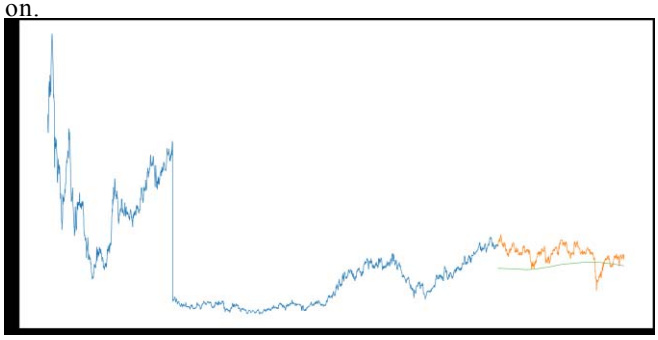


Figure 3 : Graph of predicted future when the model is built with moving average. Blue lines represent the past, yellow lines represent the actual future and green lines represents the predicted future.

The value of RMS is close to 65, r2 value is close to -12, MAE is close to 58. The results as shown in Figure 3 are not very likely along with that the expected values are similar to the values found in the trained set.

B. *Liner Regression Method*

The linear regression[2] model that presumes a relationship between the dependent and independent variables (x, y), to be linear in nature is known as linear model also termed as linear regression. The method is responsible for representing the relationship between the variables, in the form of an expression. In higher dimensions when we have more than one input (x), and they are independent of each other.

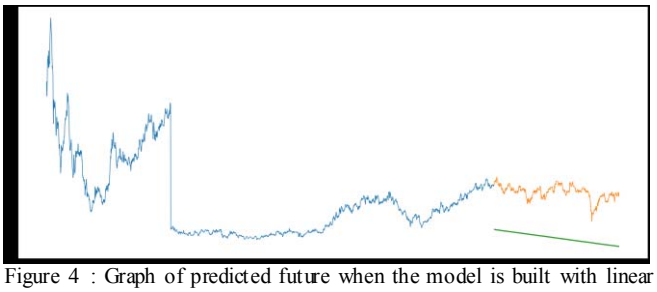


Figure 4 : Graph of predicted future when the model is built with linear regression model. Blue lines represent the past, yellow lines represent the actual future and green lines represents the predicted future.

The technique, linear regression is quite uncomplicated to elucidate. The value of RMS is around 250, r2 is around -43, and MAE is 58. Though there's a disadvantage in this technique in which the model over fits into the columns of date and month. Alternatively, the model takes the value of the same date of previous month or so or the same data of specific date and month of different year, instead of taking previous values. Thus, the model predicts wrong data.

C. *K – Nearest Neighbours (KNN) model*

In the K-Nearest Neighbors(KNN) model[3], the model tries to compare the old and new data set to set some similarity criteria. The model can be used for both regression and classification. This model assumes that similar things exist in close proximity. If k = 1, then the object or the point will simply be assigned the class nearest to it. If k = N, the object or point will be assigned to the average of N closest neighbor.

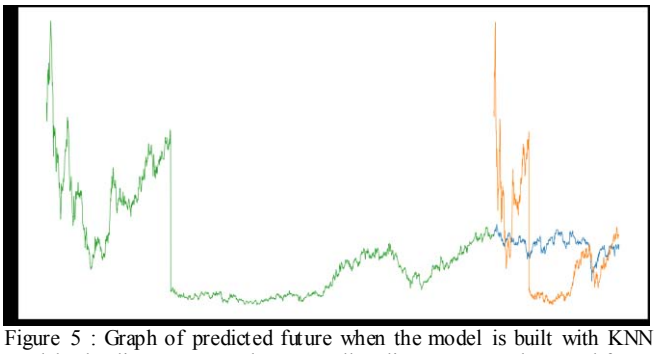


Figure 5 : Graph of predicted future when the model is built with KNN model. Blue lines represent the past, yellow lines represent the actual future and green lines represents the predicted future.

After studying the KNN model, it has been observed that the RMS value is around 210 which is quite similar to the value as calculated in the linear regression model the r2

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:34:22 UTC from IEEE Xplore. Restrictions apply.

value is -31.28 and MAE value is 170.48, along with the similarity of the plot patterns. Thus we can conclude that the regression techniques aren't operable in this case.

D. Auto Arima

For time series casting, the most in-demand statistical model is the ARIMA[4]. The former values are considered into account by the model, for the prediction of future values. The major parameters in the give model are:

- p (former values are used to predict the succeeding values)
- q (former errors in the forecast are used to forecast the hereafter values)
- d (differencing order)

The tuning of parameters for the ARIMA munches up a lot of time. Thus, auto ARIMA is used, which automatically choses the most efficient combination of the parameters (p,q,d), which result in minimal error.

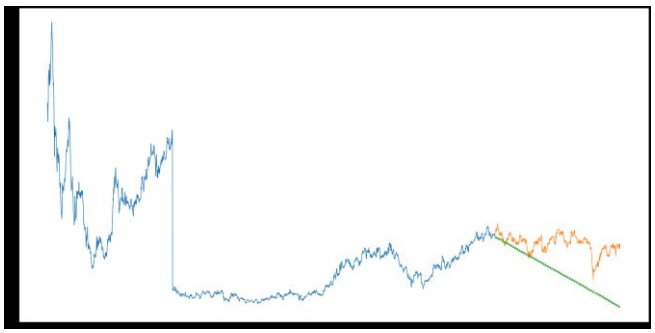


Figure 6 : Graph of predicted future when the model is built with auto ARIMA model. Blue lines represent the past, yellow lines represent the actual future and green lines represents the predicted future.

It has been seen in the past that an auto ARIMA uses the former values used for so as to understand the time series pattern. The model recorded an ascending trend within the series by making use of these values. Though the forecasting that make use of such methods are much efficient and optimized than that of the antecedently carried out models, these forecasting are yet not near to the values as expected to be in the real world. The RMS value is 143.97, r2 value is -13.81, and MAE value is 121.94.

E. Prophet

There are many time series models that could be applied to the stock price dataset, but most of them need a sizeable amount of preprocessing of data before the model can be used. The time-series library Prophet[5], which was developed by Facebook, does not require any data for preprocessing and is incredibly easy to use. The date and goal (ds and y) columns are the model's data..
In the previous data, Prophet tries to understand the trend and thus works well if the dataset is massive.

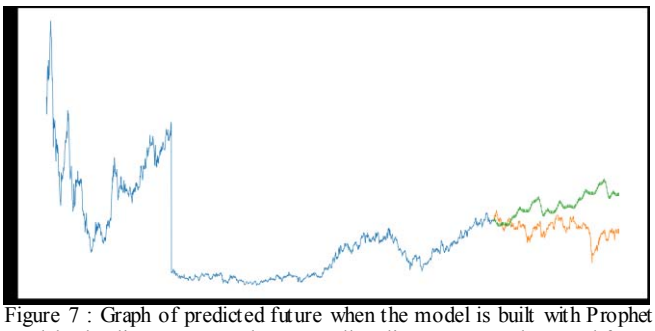


Figure 7 : Graph of predicted future when the model is built with Prophet model. Blue lines represent the past, yellow lines represent the actual future and green lines represents the predicted future.

Similar to the other time series forecasting methods, Prophet tries to record the periodicity or the trend from the formerly used data. Even though the model doesn't meet up to the results as expected in such cases, it keeps up well on the time series datasets.
The RMS value is 131.42, r2 value is -11.34, and MAE value is 109.87. As it is observed that stock prices rise and fall very frequently, depending upon what's happening in the market. Prices therefore fluctuate a lot and it becomes difficult for the forecasting methods like Prophet, ARIMA and SARIMA, to meet up with the desired output in such cases.

F. Long Short Term Memory(LSTM)

For the sequence forecasting problems, LSTMs are very commonly used and have turned out to be very successful. The explanation for the fact that they work effectively is that LSTM[5] can store the past important information and drop the information that is not. LSTM consists of 3 gates:

- The forget gate: It is used to remove the information which is not required any further.
- The input gate: The information is added by the input state to the cell state.
- The output gate: It selects the facts and figures that is to be represented as the output.

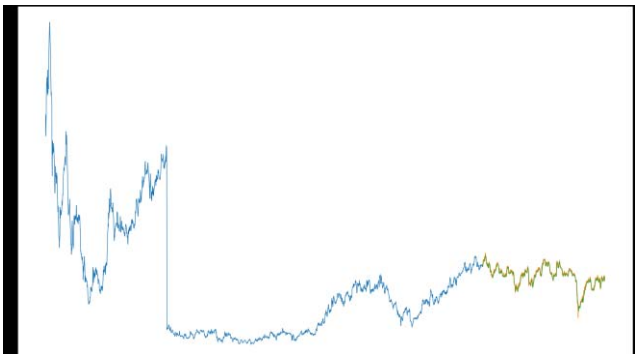


Figure 8 : Graph of predicted future when the model is built with LSTM model. Blue lines represent the past, yellow lines represent the actual future and green lines represents the predicted future.

For different parameters, such as adjusting the number of layers of LSTM, adding drop-out value, or increasing the number of epochs, the LSTM model may be modified.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:34:22 UTC from IEEE Xplore. Restrictions apply.

Year	Dataset Used	Technique	Observation	Accuracy
2019[6]	Data from internet sources related to the stocks from Istanbul.	Artificial neural network (ANN) , Artificial neuro fuzzy inference, Swarm Intelligence	• In all aspects of time, accuracy and accuracy, ANN Improved LM is far better than ANFIS and ANN traditional	ANN (Improved LM) RMSE :0.87 R^2 : 0.997 ANFIS RMSE : 1.86 ANFIS R^2 : 0.98116 ANN (Traditional) RMSE:1.881 R^2:0.98
2019[7]	Data from NASDAQ Composite (IXIC)	LSTM network Model	• The model's predictions are quite close to the real stock market behavior using historic data • Stock market cannot be determined using the historic data only but also depends on various factors other than that.	MSE : 0.0022 MAD : 0.0495
2019[8]	Stock Exchange of Thailand (SET)	CNN (Convolution Neural Network)	• CNN stock price prediction method has high accuracy and application value on the Thailand stock exchange dataset • Still need to optimize and evaluate for generic data	Represented in figures, to be estimated as 98-99%
2019[9]	February 2013 to March 2017, Standard and Poor's 500 companies	Arima, Facebook prophet, Recurrent Neural Network LSTM, Mean Absolute Percentage Error (MAPE) as evaluating method	• As opposed to standard models like ARIMA and Facebook Prophet, RNN models worked well.	ARIMA 3.89 - 17.38, For Facebook prophet 5.12 - 14.08, LSTM 1.12 - 9.8
2019[10]	January 2010 to September 2019, Shanghai Stock Exchange	LSTM	• LSTM gave promising result for the dataset used of Shanghai Stock exchange collected over a period of time.	88%-91% for each stock
2020[11]	10 year data from 4/5/2009 to 4/5/2019 of Nike, Goldman Sachs, Johnson and Johnson, Pfizer and JP Morgan Chase	Artificial Neural Network, Random Forest	• A comparison of RMSE (Root Mean Square Error), MAPE (Mean Absolute Percentage Error), and MBE (Mean Bias Error) values explicitly shows that ANN outperforms RF in terms of stock price estimation.	RMSE 0.42 - 3.30, MAPE 0.70 - 1.09, MBE = 0.0138 - 0.0762 for ANN, RMSE 0.43 - 3.40, MAPE 0.75 - 1.14, MBE 0.0138 - 0.0761 for RF
2019[12]	The Shanghai composite index, Shenzhen index, the Shanghai and Shenzhen 300 index	LSTM model	• As opposed to the actual value, the expected values high and low are postponed.	No accuracy data was provided
2020[13]	July 1, 2016 to June 30, 2018, S&P 500 index	ANN, Random Forests, Gaussian Naïve Bayes, Gradient Boosted Machines, Logistic Regression, Multinomial Naïve Bayes, Support Vector Machine, KNN, Multilayer Perceptron, Classification and Regression Tree, Linear Discriminant Analysis, AdaBoost, Extra Tree, Random Forest Classifier, Gradient Boosting Classifier.	• Gradient Boosted Machines achieves highest average accuracy of 76.50% in comparison to models like Extra Tree, random forest etc.	GBM - 76.50%, ET - 73.61%, CART - 73.27%, RF - 72.13%, LDA- 65.54, KNN - 64.88, MLP- 62.58, and SVM - 61.70%, AB - 55.85%

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:34:22 UTC from IEEE Xplore. Restrictions apply.

Year	Dataset Used	Technique	Observation	Accuracy
2018[14]	Historical data from the Taiwan stock exchange corporation (TWSE)	CRNN and LSTM model	• LSTM findings are equivalent to CRNN for a variety of stocks. • In several stock samples, the neural network consistency proposed in CRNN is more robust, and it also displays the best error rate.	No accuracy data was provided
2019[15]	Data from Inc. KAKAO from all media companies registered in NAVER	Korean Neural Language Processing	• The findings differ depending on the type of stock, natural language processing device, deep learning prediction model tool, and hyper parameter.	Average Accuracy 8-Dimensions: 0.512335 16-Dimensions: 0.513071
2019[16]	Stock market data of (a) Tesla from 29-06-2010 to 17-03- 2017 and (b) Apple from 02-01-2014 to 31-12-2018 have been used as dataset in this system.	Single level Perceptron, Multi-level Perceptron, LSTM	• After analyzing all of the methods individually, it was discovered that machine learning methods, especially MLP and LSTM, were the most effective in forecasting stock prices with the lowest MSE and MAPE values.	MAPE Values for : TESLA : SLP – 1.0094 MLP – 0.86 LSTM – 0.6993 APPLE : SLP – 0.87 MLP – 0.7907 LSTM – 0.7013
2019[17]	Stock Exchange of Thailand (SET)	CNN	• The established model can be conveniently extended to any other stock exchange with a reasonable amount of regular historical prices.	No accuracy data provided. Accuracy is depicted in form of images.
2019[18]	American airline (20 years)	LSTM with RNN by two different approach that is recursive(one value at a time) and hybrid(average of multiple prediction of same value)	• In this model the data predicted by the recursive model is more volatile than the data predicted by hybrid method.	Recursive approach MSE- 1.042e-2 Hybrid approach MSE- 5.63e-4
2019[19]	100 days of data from stock 399300,399007 and 399009	LSTM and ARIMA	• In this study author predicted the stock values for 3 different dataset with different stability that is the volatility of the stock price trend and took the average RSME for overall performance of the model	LSTM RMSE- 0.0374 ARIMA RMSE- 0.10158
2019[20]	Datasets Like Reliance, ITC, TATA MOTORS and Dr. Reddy's Lab And their respective Twitter News and Tweets regarding the stock using twitter api	Sentiment Prediction using Naïve Bayes and Multiclass SVM Stock prediction with sentiment analysis using multi-linear regression	• First predict the sentiment and the respective effect on the outcome of the future stock price using Naïve Bayes and Multiclass SVM with bag of words library and then stock prediction with sentiments analysis using Multi-linear regression	Sentiment Prediction NB Accuracy– 0.76106 MSVM Accuracy- 0.78761 Stock Prediction MLR Accuracy- 0.65 Overall MAPE – 1-5%
2019	Apple, Amazon, American Airline, Facebook, Microsoft	ARIMA, FB Prophet, Various variation of RNN LSTM that is RNN-p(RNN-LSTM with price as input), RNN-pp(RNN-LSTM with price and text polarity as input), RNN-pt(RNN-LSTM with price and text as input), RNN-mv(RNN-LSTM with multivariate model)	• Using analysis of news as sentiment analysis to study the trend of the stock and then using different method and different input types as the input and by that study found that the RNN-LSTM with price and text polarity as input gives the most accurate prediction.	MAPE ARIMA- 7.39 FB Prophet-7.98 RNN-p-2.13 RNN-pp-2.03 RNN-pt-2.17 RNN-mv-10.43

Table 1 : literature Review

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:34:22 UTC from IEEE Xplore. Restrictions apply.

IV. RESULT

In Table 2 the rows represent different techniques used like Moving Average, Linear Regression etc. and columns represent different evaluating parameters including RMS, R2, and MAE.

	RMS	R2	MAE
Moving Average	65.19206217	-2.036683033	58.03018002
Linear Regression	249.4082088	-43.43054073	247.0054433
KNN	212.5602332	-31.28301299	170.4863214
Auto Arima	143.9723997	-13.81057611	121.9446558
Facebook prophet	131.428607	-11.34212199	109.8763046
LSTM	9.424258267	0.8194021263	8.625745379

Table 2 : Results for different algorithms.

V. CONCLUSION

This paper provides a forecasting method to predict the opening prices of stocks. Based on the literature reviewed and the results encountered after running through above models on the database used, following advantages and disadvantages are observed as follows:

Moving average advantage

- Moving averages are useful for predicting goods or services with consistent demand and a moderate tendency or seasonality.

Moving average disadvantage

- Does not respond to fluctuations that occur for a cause, such as cycles and seasonal effects, as a moving average does.

Linear regression advantage

- Linear regression is easy to use and the output coefficients are simpler to analyze.

Linear regression disadvantage

- Linear regression, on the other hand, means that the dependent and independent variables have a linear relationship. That is, it is said that they have a linear relationship. It means the attributes are independent of one another.

KNN advantages

- Quick computation time
- Easy to understand the algorithm
- Versatile – can be used for regression and classification
- High precision – no need to compare to better-supervised learning models
- There are no additional assumptions about data, and there is no need to tune multiple parameters or construct a model. This makes it critical in the case of nonlinear data.

KNN disadvantages:

- Sensitive to the size of the data and irrelevant functionality
- Requires high memory – must save all of the training data
- Since it saves all of the training data, it can be computationally costly.

Facebook prophet advantages:

- While the true benefit of this model can only be evaluated on massive datasets, Prophet does allow for the forecasting of a large number and variety of time series problems — genuinely forecasting at scale.

Facebook prophet disadvantages:

- Prophet cannot account for exogenous factors that influence the time series. There is no way to describe exogenous during modeling, which puts the prophet at a disadvantage compared to ARIMA.
- Prophet is unable to accommodate multiplicative models.
- Data must be entered in a specific format.

Via a number of techniques through which we processed the stock data and reviewing past year papers mentioned in the literature review section of this paper, it can be inferred that the LSTM model produced the most positive results because it can compensate for disappearing values. However, when we take the impact of historical data into account, on the price trends it turns out that it's too singular on market patterns and might be able to predict the price on a given day completely and reliably. We may incorporate data forecasts related to stock-related news and basic details to enhance the stability and accuracy of the model in the case of a major event.

REFERENCES

[1] A. Thakur, S. Kumar, and A. Tiwari, "Hybrid model of gas price prediction using moving average and neural network," 2016, doi: 10.1109/NGCT.2015.7375218.

[2] A. Izzah, Y. A. Sari, R. Widyastuti, and T. A. Cinderatama, "Mobile app for stock prediction using Improved Multiple Linear Regression," 2018, doi: 10.1109/SIET.2017.8304126.

[3] T. Liu, J. Ma, W. Guan, Y. Song, and H. Niu, "Bus arrival time prediction based on the k-nearest neighbor method," 2012, doi: 10.1109/CSO.2012.111.

[4] P. Areekul, T. Senjyu, H. Toyama, and A. Yona, "A hybrid ARIMA and neural network model for short-term price forecasting in deregulated market," *IEEE Trans. Power Syst.*, 2010, doi: 10.1109/TPWRS.2009.2036488.

[5] C. Jeenanunta, R. Chaysiri, and L. Thong, "Stock Price Prediction with Long Short-Term Memory Recurrent Neural Network," 2018, doi: 10.1109/ICESIT-ICICTES.2018.8442069.

[6] H. N. Shah, "Prediction of Stock Market Using Artificial Intelligence," *2019 IEEE 5th Int. Conf. Converg. Technol. I2CT 2019*, pp. 1–6, 2019, doi: 10.1109/I2CT45611.2019.9033776.

[7] S. O. Ojo, P. A. Owolawi, M. Mphahlele, and J. A. Adisa, "Stock Market Behaviour Prediction using Stacked LSTM Networks∗ ," *Proc. - 2019 Int. Multidiscip. Inf. Technol. Eng. Conf. IMITEC 2019*, pp. 1–5, 2019, doi: 10.1109/IMITEC45504.2019.9015840.

[8] L. Sayavong, Z. Wu, and S. Chalita, "Research on stock price prediction method based on convolutional neural network," *Proc. - 2019 Int. Conf. Virtual Real. Intell. Syst. ICVRIS 2019*, no. Dl, pp. 173–176, 2019, doi: 10.1109/ICVRIS.2019.00050.

[9] S. Mohan, S. Mullapudi, S. Sammeta, P. Vijayvergia, and D. C. Anastasiu, "Stock price prediction using news sentiment analysis," *Proc. - 5th IEEE Int. Conf. Big Data Serv. Appl. BigDataService 2019, Work. Big Data Water*

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:34:22 UTC from IEEE Xplore. Restrictions apply.

Resour. Environ. Hydraul. Eng. Work. Medical, Heal. Using Big Data Technol., pp. 205–208, 2019, doi: 10.1109/BigDataService.2019.00035.

[10] F. Liu, X. Li, and L. Wang, "Exploring Cluster Stocks based on deep learning for Stock Prediction," *Proc. - 2019 12th Int. Symp. Comput. Intell. Des. Isc. 2019*, pp. 107–110, 2019, doi: 10.1109/ISCID.2019.10107.

[11] M. Vijh, D. Chandola, V. A. Tikkiwal, and A. Kumar, "Stock Closing Price Prediction using Machine Learning Techniques," *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 599–606, 2020, doi: 10.1016/j.procs.2020.03.326.

[12] D. Wei, "Prediction of Stock Price Based on LSTM Neural Network," *Proc. - 2019 Int. Conf. Artif. Intell. Adv. Manuf. AIAM 2019*, pp. 544–547, 2019, doi: 10.1109/AIAM48774.2019.00113.

[13] W. Khan, M. A. Ghazanfar, M. A. Azam, A. Karami, K. H. Alyoubi, and A. S. Alfakeeh, "Stock market prediction using machine learning classifiers and social media, news," *J. Ambient Intell. Humaniz. Comput.*, no. 0123456789, 2020, doi: 10.1007/s12652-020-01839-w.

[14] S. E. Gao, B. S. Lin, and C. M. Wang, "Share price trend prediction using CRNN with LSTM structure," *Proc. - 2018 Int. Symp. Comput. Consum. Control. IS3C 2018*, pp. 10–13, 2019, doi: 10.1109/IS3C.2018.00012.

[15] H. Yun, G. Sim, and J. Seok, "Stock Prices Prediction using the Title of Newspaper Articles with Korean Natural Language Processing," *1st Int. Conf. Artif. Intell. Inf. Commun. ICAIIC 2019*, pp. 19–21, 2019, doi: 10.1109/ICAIIC.2019.8668996.

[16] I. Bhattacharjee and P. Bhattacharja, "Stock Price Prediction: A Comparative Study between Traditional Statistical Approach and Machine Learning Approach," *2019 4th Int. Conf. Electr. Inf. Commun. Technol. EICT 2019*, no. December, pp. 1–6, 2019, doi: 10.1109/EICT48899.2019.9068850.

[17] A. Nakayama, K. Izumi, H. Sakaji, H. Matsushima, T. Shimada, and K. Yamada, "Short-term Stock Price Prediction by Analysis of Order Pattern Images," *CIFEr 2019 - IEEE Conf. Comput. Intell. Financ. Eng. Econ.*, pp. 1–5, 2019, doi: 10.1109/CIFEr.2019.8759057.

[18] C. Peng, Z. Yin, X. Wei, and A. Zhu, "Stock Price Prediction based on Recurrent Neural Network with Long Short-Term Memory Units," *2019 Int. Conf. Eng. Sci. Ind. Appl. ICESI 2019*, pp. 1–5, 2019, doi: 10.1109/ICESI.2019.8863005.

[19] V. Sharma, R. Khemnar, R. Kumari, and B. R. Mohan, "Time series with sentiment analysis for stock price prediction," *2019 2nd Int. Conf. Intell. Commun. Comput. Tech. ICCT 2019*, pp. 178–181, 2019, doi: 10.1109/ICCT46177.2019.8969060.

[20] F. Qian and X. Chen, "Stock prediction based on LSTM under different stability," *2019 IEEE 4th Int. Conf. Cloud Comput. Big Data Anal. ICCCBDA 2019*, pp. 483–486, 2019, doi: 10.1109/ICCCBDA.2019.8725709.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:34:22 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021)
IEEE Xplore Part Number: CFP21K74-ART; ISBN: 978-0-7381-1327-2

Adaptive Histogram Equalization based Approach for SAR Image Enhancement: A Comparative analysis

Vinay Dubey
Department of Computer Science and Engineering, Delhi Technological University
vinaydubeyamit@gmail.com

Rahul Katarya
Department of Computer Science and Engineering, Delhi Technological University
rahuldtu@gmail.com

2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) | 978-1-6654-1272-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICCS51141.2021.9432287

***Abstract*—Nowadays, flood is the most frequently occurring natural disaster among all-natural disasters. A flood is a disaster that causes huge causalities. Recovery from these huge causalities is a very serious concern for developing countries. This research work focuses on flush food, which is a very common disaster that occurs due to heavy rainfall or some manmade reasons. An Image of the flooded area provides useful information for taking necessary precautions in that particular situation. If the captured image is of poor quality, then isn't provide sufficient information. This paper considers the Synthetic Aperture Radar (SAR) images of the flooded area and it is mainly focused on increasing the image's visual quality. There are various image enhancement techniques available; this paper has used the histogram-based image quality improvement technique done by the Histogram Equalization (HE) method. To defeat the limitation of HE, another enhancement approach named Adaptive Histogram Equalization (AHE) is also used. In many cases, HE is increasing the quality of the image, but on the other side over brightness is causes the loss of some useful background information. In the AHE technique, the image is divided into small blocks: an equalization function is applied on each block. Both the techniques are compared on the various images of the flooded area, and the result shows that AHE provides better results.**

Keywords— Adaptive Histogram Equalization, Contrast enhancement, Flood management, Histogram Equalization, Image Processing, Spatial Domain

I. INTRODUCTION

The disaster causes a lot of damage and serious disruption. A disaster is an unfortunate event that causes serious disruption to human society. A natural disaster is a result of the natural imbalance that results in floods, earthquakes, etc. Flood is a very common issue in many countries. Several parts of India get highly affected by the flood. In turn, it disturbs the economic loss and human resources, affecting developing countries like India.

.The purpose of image enhancement is to extrude the important details hidden in the image and to increase the contrast of the image with less dynamic range. Several methods are available to enhance image quality, some of them working on the spatial domain and some on the frequency domain [1].

This paper focuses on enhancing the flood image to locate the flooded area in a better way. Synthetic Aperture Radar (SAR) imaging is important for Disaster, Army, and Civil applications such as target monitoring and classification. The accuracy of the application decreases with the presence of Speak and poor resolution of images created from SAR image Acquisition equipment [2]. Artificial intelligence (AI) technique for understanding remote sensing and provides very successful results. Besides, various data collecting techniques such as (SAR) are used in many applications. A satellite is an essential device that provides an image of the Earth's surface. When we talk about natural disasters such as hurricanes, tsunami, or floods, they are mostly surveillance by SAR images. [3]

- In the context of image quality improvement techniques, HE has gained the most consideration due to its smooth implementation approach and highly efficient result.
- In our paper, we use HE and the AHE techniques to enhance the SAR images of the various flooded location.
- Our implemented methods for image enhancement are efficient, and experimental results show with different images that this method is effective for image enhancement.

In the above section, we introduce our work and the remaining part of the paper contains. Section II includes the literature survey followed by section III, which includes the detailed explanation of our implemented methodologies. Section IV includes the analysis of the results of implemented techniques. Last, section V, covert the summary and conclusion of our paper; we also include future scope for getting better results.

.

II. RELATED WORK

This part of our paper is covered the previous work done to get better quality of the image. According to the paper, the author analyzes the methods for reducing noise, Particle Swarm Optimization, and mirror-extended optimization use to enhance satellite images. In this approach, the author develops a gain function that merges the feature enhancement alongside the speckle reduction [4]. In the next approach we discuss in our study, the author presents a method for enhancement of SAR image based on the Memetic Algorithm and Non-subsample Contourlet Transform (NSCT). First, a better development Function that integrates macular reduction with the characteristic

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:43:42 UTC from IEEE Xplore. Restrictions apply.

enhancement that introduces NSCT coefficient and shrinks and spreads and non-stop and multi-population co-operators is introduced to fine-tune the parameters of the enhancement function [5]. For Inverse SAR (ISAR) and SAR Images, traditional Fourier-based image reconstruction techniques provide output with limited resolution. In this paper author use modified noninteger Nyquist SVA formulation. This provided very good accuracy for radar image enhancement [6]. After more use of SAR and automatic target reorganization (ATR), these are widely used in most decision-making work. In this paper, the author focuses on the feature base noise-reducing issues for the circular SAR system. Proposed P-LSCE Able to increase target features with an Isolating speckle noise compatible with correction later verification work [7]. Among the various image enhancement methods, the histogram equalization method is commonly used because of its high accuracy and efficiency. In this paper, the author proposes a Contrast-Accumulated Histogram Equalization approach that overcomes the issue of overemphasizes and indiscriminate the contrast around intensities with a large pixel population [8]. In the case of the microscopic image, it usually contains too much noise and blurs. In this paper, the author suggested subdivision-based and fuzzy set-based image enhancement techniques. The subdivision method can conceal the noise to decompose the initial image. This is possible to reconstruct an image by splitting the low pass portion of the raw image. After the reconstruction of the image, Fuzzy Set enhances the image [9]. After the huge application of SAR images, automatic target detection (ATD) is very commonly used nowadays. In this approach, the author forces the removal of the noise from the circular SAR system. In this paper, the Lp-regularized low-rank and space-angle continuity extraction (Lp-LSCE) methods are developed by the author for reducing the noise from the reference image [10]. In the case of the 2-D, SAR image in the frequency domain with the single frequency is suffered from the depth of focus issue. This limitation leads to the difficulty of targeting the object. In this paper, the author develops a focusing-enhancement method (GAN-FEM) based on a generative adversarial network (GAN) suitable for monochromatic Synthetic Aperture Rader images [11]. SAR tomography (TomoSAR) completes 3Three-Dimensional images of the target region of interest by adding synthetic apertures for regeneration. In this paper, the author proposes a novel approach based on the Maximum-likelihood (ML) inspired approach (MARIA) to enhance the feature of the TomoSAR reconstructed image [12]. In this paper, the author uses a modified blued wheal algorithm for optimization of the blue wheal algorithm [13]. In this paper, the author focuses on improving the sharpness of the image in a virtual environment. In this paper, the Retinex algorithm is used to improve the quality [14].

III. METHODOLOGY

In image processing techniques, image quality improvement is one of the vital elements. Image enhancement is a technique to remove noise from the raw image and highlight the required information in the target image. For the enhancement purpose, Spatial Domain Methods were used and compared based on the output image in our paper.

Spatial domain Methods

In spatial domain methods, the enhancement approach is directly applied to the pixel of the image by manipulating the value stored in the pixel to get the desired result.

$$S = T(r) \quad (1)$$

In equation (1), T is the transformation that interprets pixel value in the desired pixel value S.

A. Histogram Equalization

Image enhancement is considered one of the essential sections of image processing. Its objective is to achieve intensity mapping functions such as distribution entropy the output intensity can be expanded. Although it's proper utility, HE is the completely inconsiderate intensity with a large pixel population even if they are slightly visible. In this paper, the author has proposed an approach to improve image quality using histogram equation and spatial filtering. HE is used to increasing contrast. The image extends the intensity values over the entire range. It is a commonly used image enhancement method for adjusting the contrast of the image by using its histogram [15].

In this approach, gray level transform is applied to the raw image, after that to maximize the image contrast and try to flatten the resulting histogram. In this, HE is used to enhancing the quality of medical images. In the case of Histogram Equalization, it spread the color intensity value over the full range [16]. In figure1 basic steps of HE is explained in the form of the block diagram is explained.

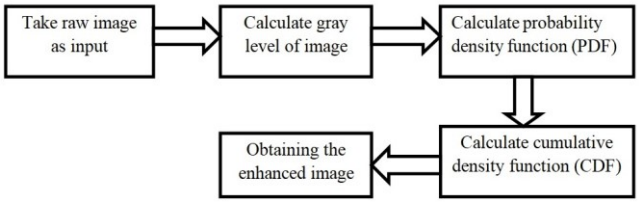


Fig. 1 Block diagram of the Histogram Equalization SAR images

a) Data collection

For improving the quality of the image for HE, we first acquire the raw image. For this, data is collected from various resources according to need. We focus on flooded areas, so we consider SAR images for the flooded area.

b) Calculate Gray level

In our approach of enhancement for HE, we apply it to the gray-scale image. For that, we convert the image to a gray-scale value. The color range of the gray-scale value is lies between min.0 (black) to the max. 255 (white).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:43:42 UTC from IEEE Xplore. Restrictions apply.

c) Calculate the PDF

Then for HE, the probability density function (PDF) is modified accordingly for the requirement. To make it easier, this technique converts the PDF of the image into equal PDF from lower pixel value to higher pixel value. When we talk about the digital image, the probability density function is a discrete function.
Let's assume a raw image (x); for example, the probability density function is calculated using probability-based on histogram $P(r_k)$ as follows. In equation (2), we show the PDF for HE.

$$PDF\ (r_k) = P(r_k) = \frac{totel\ pixels\ with\ intensity\ r_k}{totel\ pixel\ in\ image\ x} \quad (2)$$

d) Calculate the CDF

The cumulative density function (CDF) determines the amount of probability to see certain pixel intensities. From this probability density function, we calculate CDF and the intensity (r_k) of the color fluctuates between 0 (black) to (L-1) white.

$$CDF(X) = \sum_{K=0}^{L-1} P(r_k) \quad (3)$$

In equation (3) CDF is calculated where $P(r_k)$ is the probability for the pixel of intensity.
After this, we finally get the enhanced image. The final image is also on a gray scale. Invisible information also appears in the final resultant image.

B. Adaptive Histogram Equalization

When we talk about Histogram Equalization, it focuses on the global contrast of the image, so this is not efficient in all cases. In Histogram Equalization, or we can say Global Histogram Equalization, improves the contrast of the image, but it may lead to the loss of the information because it does not confine to a particular region [17]. So, in the case of the HE, this approach is efficient for image enhancement, but it may cause of loss of some useful information. AHE is widely used to overcome this issue of data loss. It is better than the ordinary histogram-based approach to improving local contrast and edges in specific areas of the image. In the following, we explain how the AHE is applied to the image.

To overcome this issue, the AHE is used.

- In this case, the image is dividing into small blocks named "Tiles," and then HE is applied to every block.
- So, in this case, the histogram is confined to small blocks. Suppose if noise is there in the block, then it will be amplified. To overcome this issue, contrast limiting is applied.
- Before applying the AHE on the gray-scale image, contras of the image are limited by applying the CDF.
- This threshold value depends on the nearby pixel or the size of the histogram of the image.
- If a histogram value is raised above the specified contras threshold, then those pixels are evenly clipped and distributed before applying histogram equalization; after equalization, bilinear interpolation is applied to remove artifacts in tile boundaries.

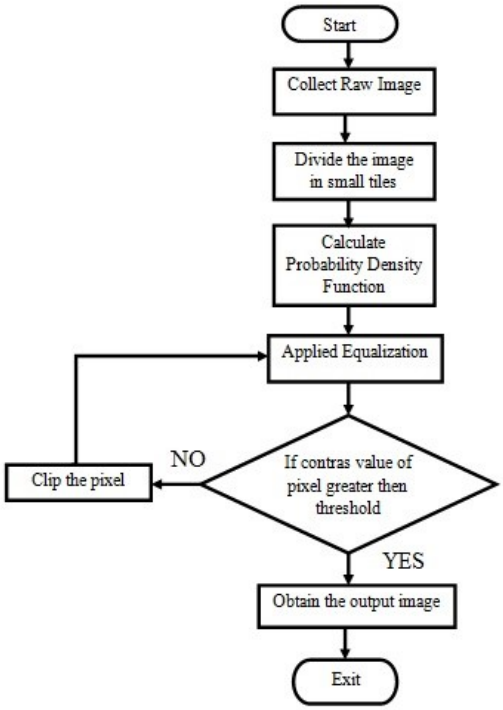


Fig. 2 Flow chart of Adaptive Histogram Equalization of SAR images

In figure 2, we explain the workflow of the adaptive histogram equalization technique. In the below flow chart, we explain the working of the AHE. The raw image is converted into tiles, and the equalization technique is applied. Based on the threshold is the condition is satisfied, we get the desired result; otherwise, clip the pixel, and then equalization is applied till getting the final result.

VI. Result And Analysis

In our paper, we have used SAR images of the various flooded area to enhance the quality of the image. For enhancement the quality, we have used HE and AHE techniques. Both the techniques are applied to the gray-scale images. The resulting final image shows that the enhanced image quality is poor after applying HE and missing some important background information. Therefore, another technique is applied to the AHE image for better results, and we obtain a better-augmented image.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:43:42 UTC from IEEE Xplore. Restrictions apply.

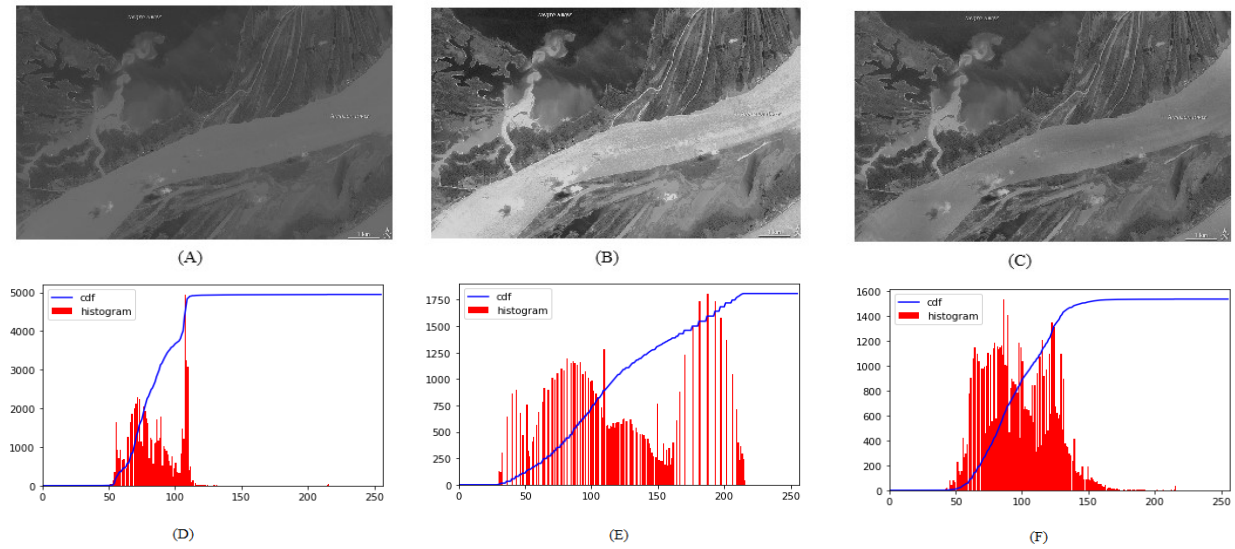


Fig. 3 (A) Raw image, (B) Enhanced by HE, (C) Enhanced by AHE, (D) Histogram of basic image (E) Histogram after HE (F) Histogram after AHE [18]

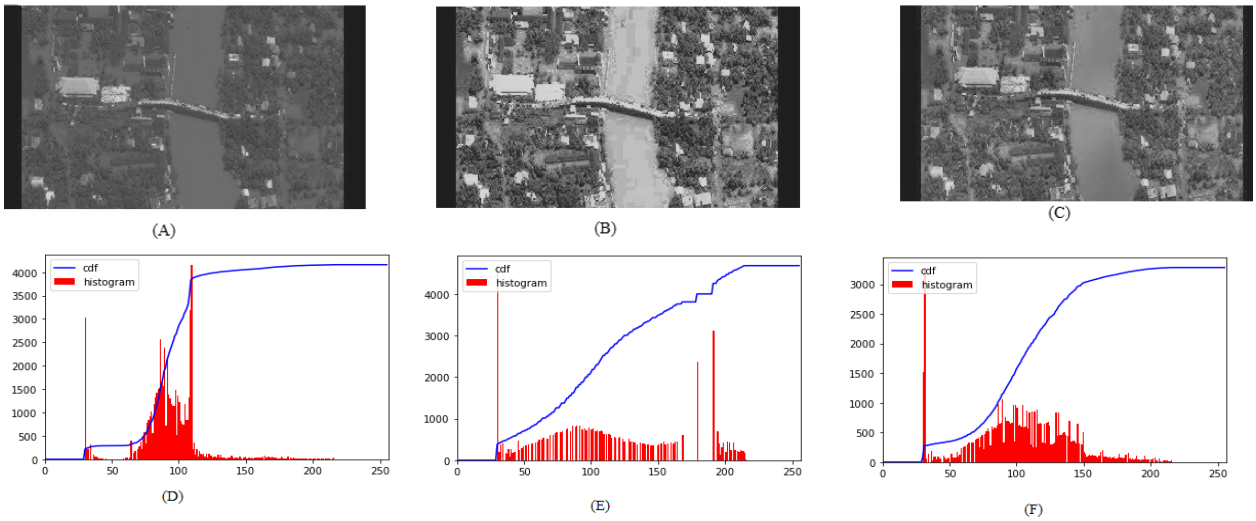


Fig. 4 (A) Raw image, (B) Enhanced by HE, (C) Enhanced by AHE, (D) Histogram of basic image (E) Histogram after HE (F) Histogram after AHE [19]

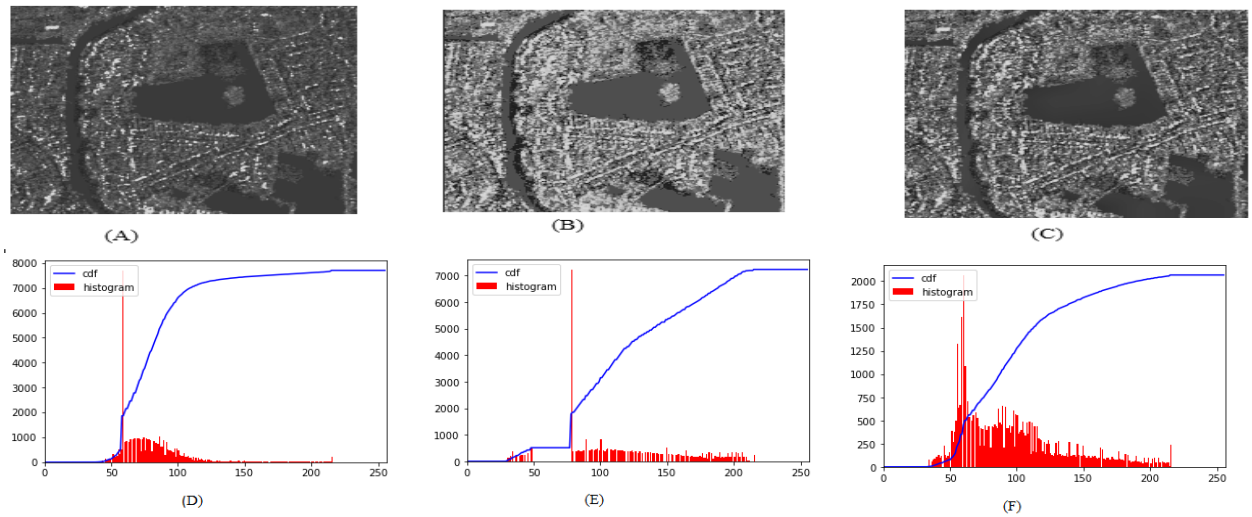


Fig. 5 (A) Raw image, (B) Enhanced by HE, (C) Enhanced by AHE, (D) Histogram of basic image (E) Histogram after HE (F) Histogram after AHE [20]

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:43:42 UTC from IEEE Xplore. Restrictions apply.

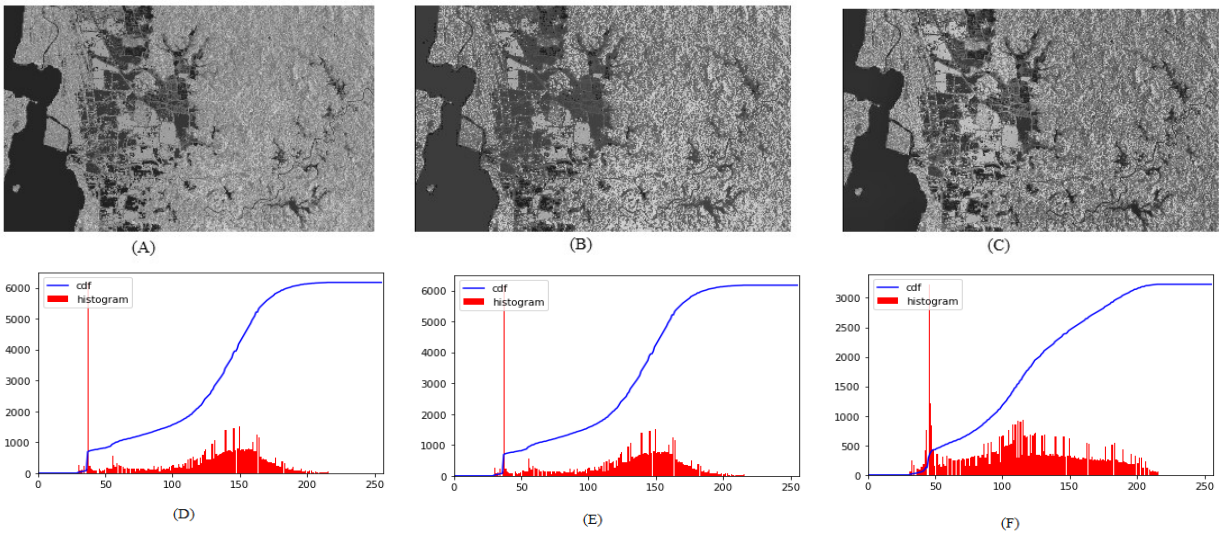


Fig. 6 (A) Raw image, (B) Enhanced by HE, (C) Enhanced by AHE, (D) Histogram of basic image (E) Histogram after HE (F) Histogram after AHE [21]

As shown in the above fig. 3, fig. 4, fig. 5, and fig. 6, we show the enhanced image by HE and the AHE techniques. In the following figure, we compared the output of our techniques. In AHE, the image is divided into tiles, and the value of the tile is (8X8) in our image.

The output of each technique is explained in the image format and also with the help of the histogram graph of each image. The graphical results of both techniques show the AHE technique performs better. The results are also compared by the histogram graph of all images. The output of the AHE technique performs better than the basic HE technique. The quality of the image is enhanced by the HE technique is loses some crucial information from the image, so to overcome this problem, AHE is applied. The output of both techniques is also visualized by the histogram graph of the output images. In the graph, the X-axis contains the gray-scale value, and Y-axis contains the pixel intensity.

V. CONCLUSION

Flood is a very commonly accruing disaster all over the world among the all-natural disasters. The reasons are natural or may be manmade. So, to reduce the casualties because of the flood, we have to detect the flooded area for detection of the flood from the SAR image. We studied and implemented the enhancement techniques on our data. In our paper, we considered the SAR image of the flooded area. The performance of the techniques is tested on the different SAR images. In this, we implemented HE and AHE on the raw SAR images of the flooded area. The output results show that the HE is enhanced the quality of the image but are lose some useful information from the image. To overcome this issue, the AHE technique is used. In many cases, HE is increasing the quality of the image, but on the other side over brightness is causes the loss of some useful background information. In AHE technique resultant image is of better quality with loss of background information. The graphical representation of the histogram of the images also proves that the Adaptive method of equalizing the histogram provides a better enhancement of the gray-scale value of the images. In the future, we can also improve the quality of the image if we limit the contras value of the image to reduce the noise. In future we enhanced the quality of SAR image using hybrid approaches.

REFERENCES

[1] F. Chen, R. Lasaponara and N. Masini, "An overview of satellite synthetic aperture radar remote sensing in archaeology: From site detection to monitoring", Journal of Cultural Heritage, vol. 23, pp. 5-11, 2017. Available: 10.1016/j.culher.2015.05.003 [Accessed 24 March 2021].

[2] W. Brown and L. Porcello, "An introduction to synthetic-aperture radar", IEEE Spectrum, vol. 6, no. 9, pp. 52-62, 1969. Available: 10.1109/mspec.1969.5213674.

[3] C. Rambour, N. Audebert, E. Koeniguer, B. Le Saux, M. Crucianu and M. Datcu, "FLOOD DETECTION IN TIME SERIES OF OPTICAL AND SAR IMAGES", ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. -2-2020, pp. 1343-1346, 2020. Available: 10.5194/isprs-archives-xliii-b2-2020-1343-2020.

[4] Y. Li, H. Gong, D. Feng and Y. Zhang, "An Adaptive Method of Speckle Reduction and Feature Enhancement for SAR Images Based on Curvelet Transform and Particle Swarm Optimization", IEEE Transactions on Geoscience and Remote Sensing, vol. 49, no. 8, pp. 3105-3116, 2011. Available: 10.1109/tgrs.2011.2121072.

[5] Y. Li, J. Hu and Y. Jia, "Automatic SAR image enhancement based on nonsubsampled contourlet transform and memetic algorithm", Neurocomputing, vol. 134, pp. 70-78, 2014. Available: 10.1016/j.neucom.2013.03.068.

[6] Xiaojian Xu and R. Narayanan, "Enhanced resolution in SAR/ISAR imaging using iterative sidelobe apodization", IEEE Transactions on Image Processing, vol. 14, no. 4, pp. 537-547, 2005. Available: 10.1109/tip.2004.841198.

[7] L. Chen, X. Jiang, Z. Li, X. Liu and Z. Zhou, "Feature-Enhanced Speckle Reduction via Low-Rank and Space-Angle Continuity for Circular SAR Target Recognition", IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 11, pp. 7734-7752, 2020. Available: 10.1109/tgrs.2020.2983420.

[8] X. Wu, X. Liu, K. Hiramatsu and K. Kashino, "Contrast-accumulated histogram equalization for image enhancement", 2017 IEEE International Conference on Image Processing (ICIP), 2017. Available: 10.1109/icip.2017.8296871 [Accessed 24 March 2021].

[9] Guo Xian Jiu, Jiang Feng Jiao and Li Xiang, "Image enhancement method based on fuzzy set and subdivision", 2011 3rd International

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:43:42 UTC from IEEE Xplore. Restrictions apply.

Conference on Awareness Science and Technology (iCAST), 2011. Available: 10.1109/icawst.2011.6163135 [Accessed 24 March 2021].

[10] L. Chen, X. Jiang, Z. Li, X. Liu and Z. Zhou, "Feature-Enhanced Speckle Reduction via Low-Rank and Space-Angle Continuity for Circular SAR Target Recognition", IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 11, pp. 7734-7752, 2020. Available: 10.1109/tgrs.2020.2983420.

[11] G. Ye, Z. Zhang, L. Ding, Y. Li and Y. Zhu, "GAN-Based Focusing-Enhancement Method for Monochromatic Synthetic Aperture Imaging", IEEE Sensors Journal, vol. 20, no. 19, pp. 11484-11489, 2020. Available: 10.1109/jsen.2020.2996656.

[12] G. Martin del Campo, M. Nannini and A. Reigber, "Towards Feature Enhanced SAR Tomography: A Maximum-Likelihood Inspired Approach", IEEE Geoscience and Remote Sensing Letters, vol. 15, no. 11, pp. 1730-1734, 2018. Available: 10.1109/lgrs.2018.2858571.

[13] S. Dutta and A. Banerjee, "Highly Precise Modified Blue Whale Method Framed by Blending Bat and Local Search Algorithm for the Optimality of Image Fusion Algorithm", December 2020, vol. 2, no. 4, pp. 195-208, 2020. Available: 10.36548/jscp.2020.4.001 [Accessed 16 April 2021].

[14] D. R., "Improved Image Processing Techniques for User Immersion Problem Alleviation in Virtual Reality Environments", Journal of Innovative Image Processing, vol. 2, no. 2, pp. 77-84, 2020. Available: 10.36548/jiip.2020.2.002 [Accessed 16 April 2021].

[15] U. Acharya and S. Kumar, "Genetic algorithm based adaptive histogram equalization (GAAHE) technique for medical image enhancement", Optik, vol. 230, p. 166273, 2021. Available: 10.1016/j.ijleo.2021.166273.

[16] M. F .Abubakar," Image Enhancement using Histogram Equalization and Spatial Filtering". International Journal of Science and Research (IJSR), India Online ISSN: 2319-7064, Volume 1 Issue 3, December 2012.

[17] M. Hitam, W. Yussof, E. Awalludin and Z. Bachok, "Mixture contrast limited adaptive histogram equalization for underwater image enhancement", 2013 International Conference on Computer Applications Technology (ICCAT), 2013. Available: 10.1109/iccat.2013.6522017 [Accessed 24 March 2021].

[18] "Flooding near Manaus, Brazil", Earthobservatory.nasa.gov, 2021. [Online]. Available: https://earthobservatory.nasa.gov/images/39359/flooding-near-manaus-brazil. [Accessed: 24- Mar- 2021]

[19] "Satellite images show devastating Kerala floods", Geospatial World, 2021. [Online]. Available: https://www.geospatialworld.net/videos/satellite-images-show-devastating-kerala-floods/. [Accessed: 24- Mar- 2021]

[20] D. Mason, S. Dance, S. Vetra-Carvalho and H. Cloke, "Robust algorithm for detecting floodwater in urban areas using synthetic aperture radar images", Journal of Applied Remote Sensing, vol. 12, no. 04, p. 1, 2018. Available: 10.1117/1.jrs.12.045011.

[21] "Monitoring monsoon flooding in Kerala with SAR and AI | Planet Watchers", Planet Watchers, 2021. [Online]. Available: https://www.planetwatchers.com/monitoring-monsoon-flooding-in-kerala-with-sar-and-ai/. [Accessed: 24- Mar- 2021]

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:43:42 UTC from IEEE Xplore. Restrictions apply.

Turkish Journal of Computer and Mathematics Education Vol.12 No.13 (2021), 1761-1765

Research Article

ADAS Authentic Data Allowed Security In Internet Of Things

Raveena Yadav [a], Vinod Kumar[b]

[a] Research Scholar, Department of Computer Science Engineering, Delhi Technological University, Delhi, India
[b]Associate Professor, Department of Computer Science Engineering, Delhi Technological University, Delhi, India

Article History: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 4 June 2021

Abstract: With the growing of population, things and networks, things connectivity is also taking a graph in exponential manner. This whole connectivity makes a whole network that is known as Internet of Things. But these devices are prone to cyber-attacks, so, security is one of the apex concerns and challenge for internet of things. Main motivation of writing this paper is that security is the main concern whether it is internet or internet of things. In this paper, we have proposed a model Authentic Data Allowed Security, in which only authentic packets are allowed to enter and usable by node and all other packets will be dropped by the router

Keywords: Internet of Things, Security, Authentic data

1. Introduction

Internet is one of the networks, in which our confidential data move from one device to another. So, there is a chance that intruder is waiting for accessing the data. That's why cyber-attack is increasing day by day, besides these new measures are also taken up for nurturing the privacy of data. In this high-tech world, devices are connected to internet and doing communication without the interference of human being and transferring the data from one node to another node and this is we can say Internet of Things (IoT). The idea of connecting devices to the internet, has come in the 90's and the term of IoT is coined by Kevin Ashton in 1999. There are many definitions given by different scientist and researchers. In [2], author described IoT as a revolutionary and global network, in which devices are getting smarter day by day and do the communication. In [5], IoT as the intelligent network in which smart devices can be tracked, monitored and managed remotely. IoT can be defined as the virtual connection between the two or more than two things and these connections are maintained by different networks such as WiFi, Bluetooth, Zig bee, RFID and many more networks are presents as shown in Figure 1.

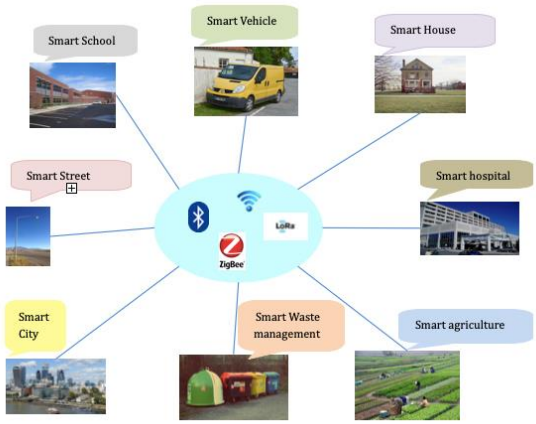


Figure 1: IoT in different sectors

There are several types of security issues happened and still happening in the IoT world such as intruder can access the information and misuse data, send the data continuously so that node or device would not able to use the resources, transfer the data to the wrong intruder and many other ways attacks can happen.

IoT can be understood with a contrast and comparative relation with the human body as shown in Figure 2. Human body has five sense organs; eyes, ear, nose, tongue and skin like this IoT has sensors that help in sensing the change in surrounding so that actuators can perform action. Nerves and neurons in the human body help in transmitting the messages, blood and oxygen, IoT has its different networks Bluetooth, Wi Fi, LoRa, Zig bee and many more through which communication happens in the IoT network.

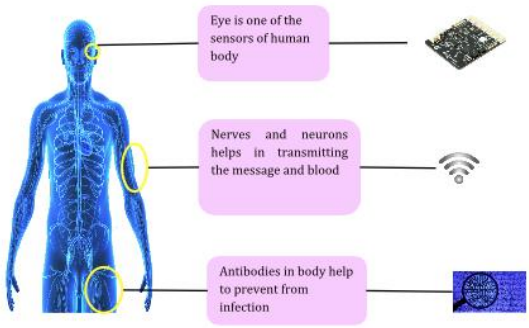


Figure 2 : IoT Human body contrasting with

This paper is divided into numerous sections, in the first section we already have discussed the Introduction of Internet of Things, in second section different layer of IoT, in section third various issues in IoT, in fourth section proposed model and in fifth section described the future scope of this research.

2. Layers in IoT

For better understanding of the complex network of IoT, researchers had divided the network into three layers, four layers and five layers. In this section, we discussed three layered model of architecture as shown in figure 3.

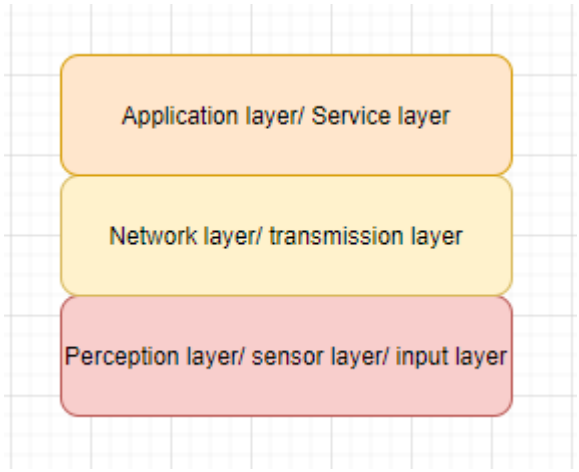


Figure 3: Three-Layered architecture of Internet of Things

Input layer: Perception layer of the IoT architecture takes the input from its surrounding. This layer can be called as sensor layer or input layer or perception layer. The main component of this layer is sensor, RFID and actuators, that helps in getting the information for further processing and identifying the objects uniquely. The range of sensor depends upon the power of sensor and area covered by that sensor.

2.1..Network layer

This layer helps in transmitting the message from one device to another device. In comparison to the human body like blood helps in transmitting the nutrients from one organ to another organ, network layer helps in sending the data. For transmitting the message network layer use WiFi, bluetooth, Zigbee, LoRa WAN and others technology.

2.2..Application Layer

Application layer on which IoT application are build and helps in managing the applications of IoT. This layer uses many protocols CoAP (Constrained Application Protocol), MQTT (Message Queue Telemetry Transport), XMPP (Extensible Messaging and Presence Protocol), AMQP (Advanced Message Queuing Protocol) to manage the application. These protocols are made for the constrained environment and act like Publish and Subscriber

3..Issues in IoT

The trending topic, IoT has given us a lot of advantages with this there are some issues also present, in this section we are going to discuss about some of the main issues.

3.1.Hardware issue

Hardware issue in IoT is basic concern, in which manufactures can left the loop hole unintentionally, this can leak our confidential data.

One more thing is related to hardware issue is that any customer buys an IoT device and try to find the flaws in device so that he can attack on other devices.

3.2.Security

Security is one of the major issues in IoT. IoT network contains a lot of data and that includes our personal and confidential information. It is one of thing that has not achieved 100 percent till date. There are some customers who have trust issue in using the devices that are connected to internet (IoT device).

a) This issue can be solved by using the encryption and decryption.

b) By using an authentication mechanism before accepting the data.

3.3.Energy management

In the perception layer, sensors take the information from surrounding. Sensors are low powered device if these will not use properly then probability of dying is very high. So, there should be a proper management of sensors. Researchers have given many algorithms for efficient management of wireless sensor network like LEACH, TEEN, PEGSIS and many others.

3.4.Data storage

As day by day, IoT network is growing with a very high pace. IoT things generate a huge amount of data in zettabyte. Maintaining and analyzing this heavy amount of data is one of the issues. This issue can be solved by two ways;

a) By removing the redundant data.

b) By storing only useful and informative data.

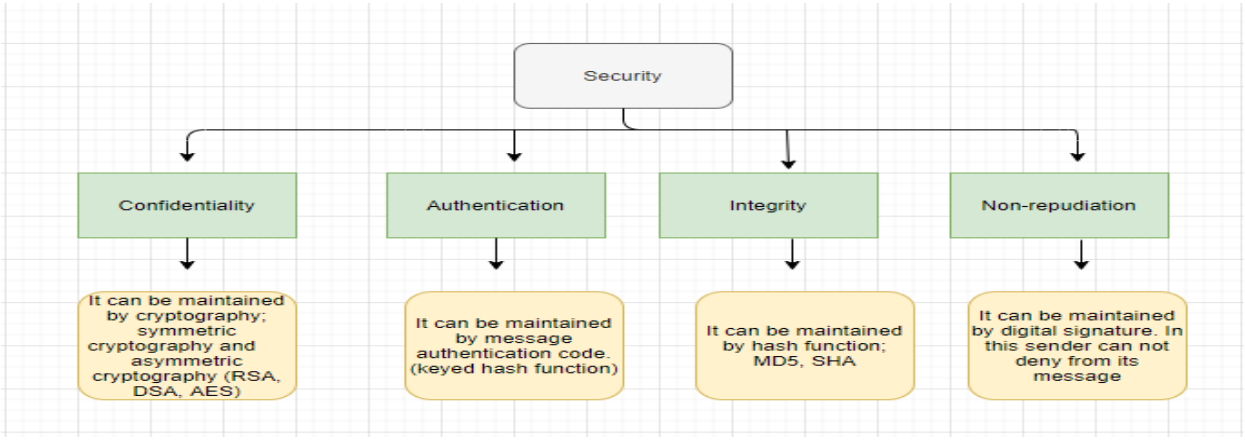


Figure 4: Types of security issues

4. Proposed Model

In proposed model, an authentic model for IoT device has introduced in which each IoT device has its own cloud storage in which all authentic IPs store. In this, a list of authentic IPs maintains and an intruder list in which all unauthentic IPs will store, it also updates data dynamically and set a threshold value.

When an IoT device receives a message from an authentic device, it will accept the packet and process it, if packet's IP neither belong to authentic IP nor intruder list then it will check its payload value, if its payload value is greater than threshold value then packet will be discarded and its IP address stores into intruder list else it will be accepted and added into authentic list if it receives from an unauthentic device then it will add it into the intruder list as shown in Figure 5.

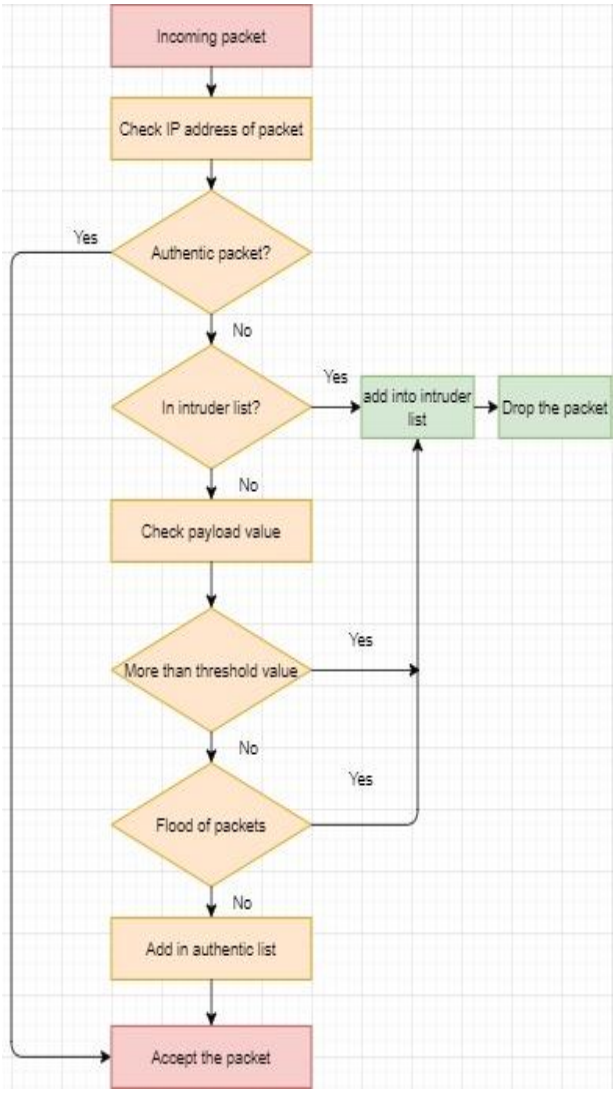


Figure 5 : Proposed framework

5. Future scope

Security is the major problem in internet world. One of the things of internet is internet of things that contains a huge amount of data, for securing this data, security is very important. In this we have a model for the security that basic function is accepting only authentic packet. The future scope of our research is implementation of our proposed model with the efficient energy management of sensor network. This combination will help in maintaining the perception layer of IoT.

6.Conclusion

With the increase of number of connected devices in IoT network, security issues and number of complaints regarding security will also increase. Both are these directly proportional to each other. In our research paper, we have mentioned about the basic information of Internet of things. There is no standardization of architecture for IoT, in this paper we have mentioned the three-layered architecture (Input layer, transmission layer and service layer) and some issues that can be happened in IoT. One security model related to authentication has been introduced in this paper. That allow only authentic packet to accept.

References

1. Ahmad W. Atamli and Andrew Martin,"Threat-based security analysis for the internet of things," in International Workshop on Secure Internet of Things. IEEE, 2014.
2. Michael Schiefer, "Smart home definition and security threats," in International Conference on IT Security Incident Management & IT Forensics. IEEE, 2015.
3. Talwana Jonathan and Huang Jian,"Smart world of internet of things and it's security concern," in IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and

Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData). IEEE, 2016.

4. Shao Xiwen, "Study on Security issue of internet of things based on RFID," in International Conference on Computational and Information Sciences. IEEE, 2012.
5. Alexandru Lavric andValentin Popa, "Internet of things and LoRa low power wide area networks challenges," in International Conference on Electronics, Computers and Artificial Intelligence (ECAI). IEEE, 2017.
6. Subha Koley and Prasun Ghosal, "Addressing hardware security challenges in internet of things: Recent trends and possible solutions," in International Conference on Ubiquitous Intelligence and Computing. IEEE, 2015.
7. Jean Pierre Nzabahimana, "Analysis of security and privacy challenges in internet of things," in International Conference on Dependable Systems, Services and Technologies (DESSERT). IEEE, 2018.
8. Anna Katrina Gomez and Simi Bajaj, "Challenges of testing complex internet of things devices and system" in International Conference on Knowledge and Systems Engineering (KSE). IEEE, 2019.
9. Xuemei Li and Li Da Xu, "A review of internet of things- resource allocation," in IEEE Internet of Things Journal. IEEE,2020.
10. Chang Le Zhong, Zhen Zhu and Ren Gen Huang, "Study on the IoT architecture and Access Technology," in International Symposium on Distributed Computing and Applications to Business, Engineering and Science (DCABES). IEEE, 2017

Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021)
IEEE Xplore Part Number: CFP21K74-ART; ISBN: 978-0-7381-1327-2

AI based Impact of COVID 19 on food industry and technological approach to mitigate

1st Ram Murti Rawat
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
rammurtirawat@dtu.ac.in

3rd Ankit Jose Toppo
Department of Computer Science & Engineering
Delhi, Technological University
Delhi, India
ankit.kdg3@gmail.com

2nd Ashish Rana
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
mrashishrana8@gmail.com

4th Akshit Beck
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
akbeck.beck7@gmail.com

2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) | 978-1-6654-1272-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICCS51141.2021.9432152

***Abstract*—The food industry or the restaurant business has always been one of the most profitable and growing businesses. As technology is evolving day by day and to sustain itself in the food industry the restaurant needs to come up with new and innovative services which they can provide to the customers. Pandemic has made a major impact on the business of the restaurant industry in 2020. This paper will help to understand how technology can help in food ordering, while taking care of the COVID-19 pandemic guidelines. The purpose of this paper is to analyze the impact of COVID-19 in the food industry and suggest methods that would help restaurants to adapt to challenges that originate from COVID-19. One of the ways would be to automate the food ordering process. Our Project "Contactless food ordering system" is a mobile application operated by a voice assistant. That the client (The Restaurant Customer) can use to scan through the restaurant menu and view various food categories and items like "Starter" "Drinks" etc. and place their order through the Application. They can modify and confirm their order, as well as make their payment. The aim of the study is to reduce the point of contact during COVID-19 and automate the ordering and billing process, simplifying the work. Thus, providing a totally computerized, automated, and scalable food ordering system that will assist the business by reducing their workload, provide better management and smooth operation. The restaurant can also use various other means to increase their presence as a business.**

Keywords— Food Industry, COVID-19, Automated System, Digital Marketing, Artificial intelligence, Voice assistant, Supply Chain

I. INTRODUCTION

Today the food industry is a multi-billion dollar economy, with a growth rate of more than 15% than any other industry. But the foodservice industry has not been able to keep up with this boom. The quality of service has declined. You often hear horror stories, which go along the lines of rude service, long waiting periods, mixed-up orders, etc. The quality of service has declined in order to increase profit.

The main reason behind this, is the lack of proper infrastructure, overburdened workforce, small profit margins, and cutthroat competition. A big food chain like KFC and Dominos are large enough that a few inconveniences do not make a big bump in their reputation, but for an individual/ independent restaurant, their reputation is everything as customer loyalty and customer retention is very important for them.

A good chunk of waiters/server's time is used up while taking orders. They have to memorize the order, as well as remember who placed the order. They have to provide recommendations, advice and deal with customers if any food items go out of stock. And let's not forget the curveball the customers throw as they change tables for a better view or better spacing. Let's just say that a waiter/server has their hands more than full. Also, let's not forget how the COVID-19 regulations have made things even more difficult for the food industry.

The customer base for these restaurants have decreased significantly due to COVID-19 restrictions.

The main goal of our project is to implement an infrastructure through our project that will help the restaurant manage the workforce and simplify the order taking/placing, payment process and reduce point of contact between customer and workers to ensure the highest safety standards. This will also reduce the workload for its workers.

In our ideal use case, each table will be equipped with a tablet via which the customers can place and manage orders via voice assistant, they will also be provided with a cashless payment option. The orders will be directly sent to the kitchen with all the relevant details. All the tablets and backend servers will be connected via Wi-Fi. Since the entire setup is offline. There is no cost of hosting an online server [3].

Small scale Restaurants and cafes will benefit from our application as it will help them do their functionalities more accurately and quickly. The restaurant can change its menu in real-time if a food item goes out of stock. Update their menu effortlessly etc.

Currently the 21st century is going through a technological boom. The Artificial Intelligence technology has advanced significantly in the past few years. The roles and

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:38:49 UTC from IEEE Xplore. Restrictions apply.

applicability of AI techniques has been continuously changing. From being applied in self-driving technology to predicting weather patterns. It has advanced to the point where AI can be used for real time anomaly detection in a system and thus prevent malware or unauthorized access to the system. The tech community is finding innovative ways to productively implement AI for betterment of the human race[8,10]. Artificial intelligence-based voice recognition technology is rapidly changing, adapting and continues to be applied to daily use products. Lights, fridges, washing machines and even home security systems are being integrated with voice assistants like Alexa or Siri. So, using voice assistant to place orders at a restaurant seems natural [9].

II. LITERARY SURVEY

Early estimates predicted that most major economies would lose 2.9 % of their GDP if COVID-19 coronavirus became a pandemic. This forecast has already been revised to a GDP loss of 4.5%. This economic damage is largely due to the fall in demand, meaning less consumers are willing to purchase goods and services available in the economy [11].

Restaurants are shutting down, as they are unable to generate profit. Even takeout and delivery orders are unable to generate enough profit to make up for the lost dine-in sales. According to USB Evidence Lab restaurant's revenue remained in the red as dine-in sales plunged 69% as the week ended on November 29, 2020. Despite the fact that takeout and delivery orders increased 59%. Industry experts predict that the winter season would further increase the difficulties of the restaurants during the COVID-19 pandemic. Harsh winters would mean even lesser customers that willing to opt for the dine-in experience [4].

Voice recognition has improved significantly in recent years and has made several waves of major innovation. Voice commands for search, dictation have become a standard feature on a smart device such as a smartwatch, Alexa mini, etc. Voice technology has advanced to the point that it has become one of the cornerstones of consumer's search activities, with an estimated one billion voice searches performed each month. Experts have predicted that nearly half of all searches will be voice-based in a couple of years.

Consumers are most likely to use voice technology to seek out restaurant information, with 53 percent indicating they would like to use such solutions to search for restaurants based on menu information. Forty-nine percent use voice offerings to find information about specific restaurants, and 44 percent use it to find restaurants based on their cravings.

Research shows that consumers are most likely to use voice search to find out about restaurants, with 53% indicating that they would likely use such a solution for eateries and restaurants based on menu information. With 44% use voice solutions to find out about specific restaurants and 49% use voice solutions to find restaurants based on a specific food item or their craving.

An automated food ordering system that keeps track of user's orders can be deployed by the means of an android application on which can be accessed via Voice Assistant or touch on a tablet. This system can be implemented using open source code for JAVA and android for the front end and MySQL database for the backend [5].

A. *Benefits of Voice technology during COVID-19*

The majority of the restaurants and fast-food joints operate out of rented real-estate and thus pay huge amounts of rents. COVID-19 pandemic has made it tough for businesses to pay this rent due to decreased footfall and decreased revenue.
Therefore business owners need to adapt and integrate in-demand technologies. Voice technology can help businesses and restaurants in the following ways

Voice-powered kiosks can help customers to self-order and self-pay at limited-service restaurants. Where there is no space for human contact, Alexa's Drop-In feature helps users to check the health of people in isolation.

These voice skills can provide increased customer engagement and brand experience. Get more customers loyal to your brand, in the times where all are crying for this to happen.

Make huge savings by bringing automation in the tasks where there is no need for human involvement [6].

B. *Impact of COVID-19 on businesses*

Despite the extraordinary efforts and measures by government to provide relief small businesses and supports firms, some countries are expected to see a record high bankruptcy filing (see Fig.1) [7].

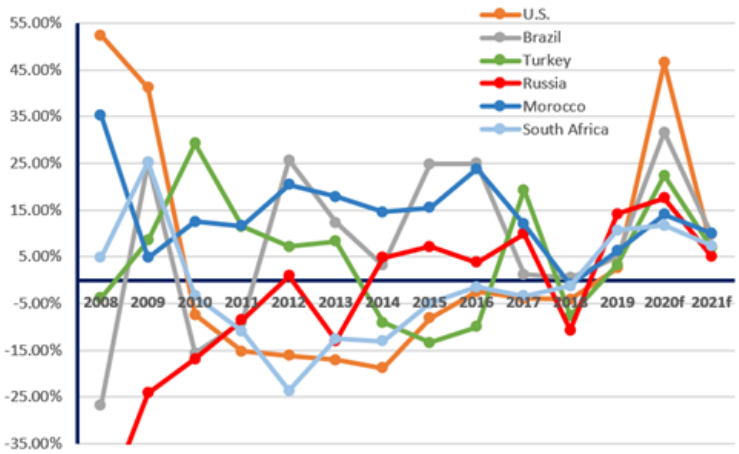


Fig. 1. Evolution and forecast of bankruptcy filings (2008–21) (percentage increase year-over-year)

Global food prices increased up to 20% during the last year (Jan 2020-Jan 2021), this increase in price was consistent with movement of other commodity prices and US currency trends [1].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:38:49 UTC from IEEE Xplore. Restrictions apply.

Due to COVID-19, big restaurant chains that have high revenue and good number of reserves can easily weather the storm. Whereas the same could not be said about the small, independent restaurants. Fate of these restaurants is one of the most debated subjects in the United States. The U.S has around 370,000 independent restaurants which make up for 57% of that total restaurants, most of which fall under on dine-in service category. Experts estimate that around 5% to 30% of the restaurants might permanently close down. Which would result in a loss of around 20,000 to 110,000 venues [2].

"Assuming 10% to15% of independents close, we estimate this would put about $20 billion to $35 billion in sales up for grabs, mostly in the full-service category. In fast food, using those same assumptions, we estimate that 2% to 4% of industry sales would be freed up," says Glass.

C. Contactless Payment Technology

Contactless payment technology has increased in popularity recently, as contactless payment has been realized as one of the methods to reduce the risk of COVID-19 transmission. Contactless payment can be described as any payment transaction that requires and absence of physical contact between the consumer's payment device and the physical terminal. The difference between "contactless payment" and "remote payment" is that remote payment makes use of internet connection via cellular network or Wi-Fi during transactions while contactless payment generally uses near field communication (NFC) technology, which does not require internet connection [12].

D. Impact of COVID-19 on Supply Chain

The food supply chain can be divided into 5 stages i.e. agricultural production, post-harvest handling, food processing, distribution and consumption.

In the earlier stages of pandemic, the supply chain was disrupted due to panic buying and hoarding. The supply could not keep up with the demand.

Government introduced safety measures to ensure the continuity of food flow in each stage. Retail workers, logistic workers etc. were deemed essential and provided exemptions from lockdowns. Proper PPE kits were provided, and social distancing guidelines were introduced as last stages of food supply chain can be severely affected by COVID-19 [13].

Due to significant restrictions in transport of goods due to the pandemic, as well as in migration of labor. Reports show that in countries such as France food distribution declined as much as 60% since restrictions which was 30% before the pandemic [14, 15].

Close-down of food processing plants also created a ripple effect in the food supply chain. Increased demand resulted in empty shells and decreased supply cased increase in price of non-regulated products such as eggs, meat products. Many farm producers were forced to cull animals due to increasing losses due to disrupted supply chain [16].

III. PROPOSED SYSTEM

Our research work's aim is to analyse that market and suggest changes, design, and develop a cheap and easily scalable Wireless food ordering system that replaces the traditional ordering system. This new system will help the restaurant function fluently with a minimum workforce while following COVID-19 norms. This new system will provide better overall customer service by reducing human error, by improving management aspect and being faster.

In our ideal use case, each table will be equipped with a tablet via which the customers can place and manage orders via voice commands or manually, they will also be provided with a cashless payment option. The orders will be directly sent to the kitchen with all the relevant details. All the tablets and backend servers will be connected via Wi-Fi. Since the entire setup is offline. There is no cost of hosting an online server.

Small scale Restaurants and cafes will benefit from our application as it will help them do their functionalities more accurately and quickly. The restaurant can change its menu in real-time if a food item goes out of stock. Update their menu effortlessly etc.

Due to COVID-19 Pandemic, orders in restaurants' have not only reduced but also shifted from dine-in to mostly take-out. As mentioned earlier customers are more likely to search for individual restaurants or search for their favourite dish online and so on. With profits taking a huge dip due to the pandemic, independent restaurants, need to do everything in their power to attract more customers. Here are some of the ways restaurants can do that.

Set up Google My Business – Independent restaurants need to make their business visible to the internet. Setting up a Google My Business helps with that. Your Google My Business should have all the relevant information like proper contact info, website, working hours with days off along with official social media handles.

Social Media Marketing / Digital marketing – Independent Restaurants cannot rely on their usual ways of promotion like word of mouth or advertisement on newspaper/magazines or billboards and with decreased revenue, setting aside capital specifically for advertisement would hurt the profitability of the restaurant even more. Restaurants should use social media like Instagram or Facebook to promote their restaurant.

Maintain the quality standard – Restaurants need to maintain their quality standard. If the quality of the food decreases, their existing customer base might also diminish. Restaurants need to make sure that their food is presentable and passes the quality check every time. There are a lot of opportunities for a restaurant to impress a customer during a dine-in experience. There are a lot of factors such as ambience, waiting period, service quality, etc. but with a take-out or a delivery, the most important factor is food

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:38:49 UTC from IEEE Xplore. Restrictions apply.

quality. So a restaurant should make sure that it doesn't decline.

Partner with a local delivery partner – Since Take-out and food deliveries have skyrocketed during the COVID-19 pandemic. Restaurants should consider partnering up with their local food delivery partner to increase their restaurants' reach. Restaurants should also consider taking orders directly through a website or a number to avoid giving a cut to the delivery partner .

IV. SYSTEM ARCHITECTURE

The architectural design consists of three main users: - Service Consumer, Owner of Mess/Restaurant/Manager, and Employee of mess(Chef/ Waiter).

The customer/consumer enters the restaurant, takes a table. Each table is fixed with an android tablet. The device contains the food ordering app that helps customers place the order. The application can be operated using a voice assistant or using touch inputs. The customer can use the voice assistant feature to order the food. The virtual assistance helps the customer to use the services of the restaurant while following all the COVID 19 norms.

The placed order is received by the kitchen and the chef starts working on the order. After the order is prepared, it is served by the waiter and final bill is generated is displayed in the tablet.

The customer can pay digitally using the options provided in the tablet or via cash.

The manager can change menu in real-time depending on the food stock available and view transactions at the end of day.

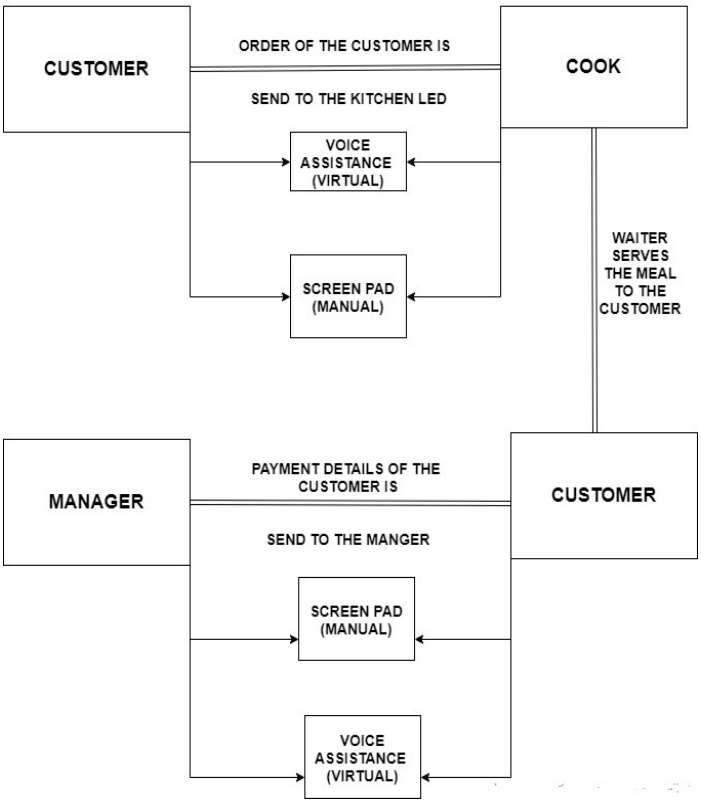


Fig. 2. System Architecture for the food ordering application.

V. ALGORITHM

A. Food Ordering App Algorithm

1) The customer is welcomed with a greeting message displaying on the tablet screen on their table.
2) He can then browse through and choose to avail of the dishes.
3) The customer can move in the menu and lock a dish, without touching the screen. Customers order his food by voice command. After selecting he can confirm the order.
4) Then the ordered details go to the kitchen. Confirmation of the order is shown.
5) Then the chef along with the kitchen staff prepares the dish and then the waiter takes it to the table which orders that dish.
6) The customer after finishing the food can pay through the 'Payment Portal'. This portal has two modes of payment 1. Cash, 2. Online Transaction or Card.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:38:49 UTC from IEEE Xplore. Restrictions apply.

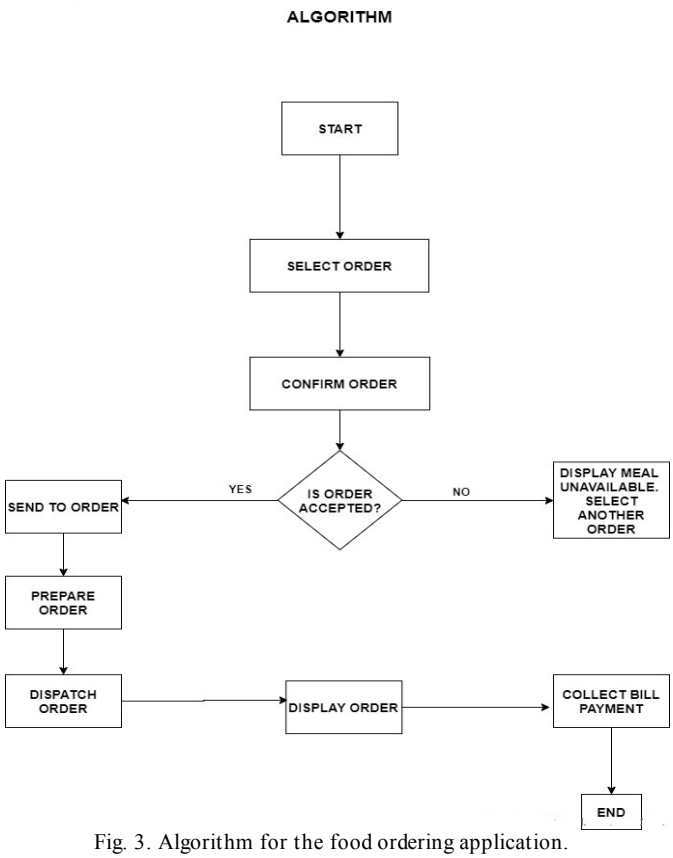


Fig. 3. Algorithm for the food ordering application.

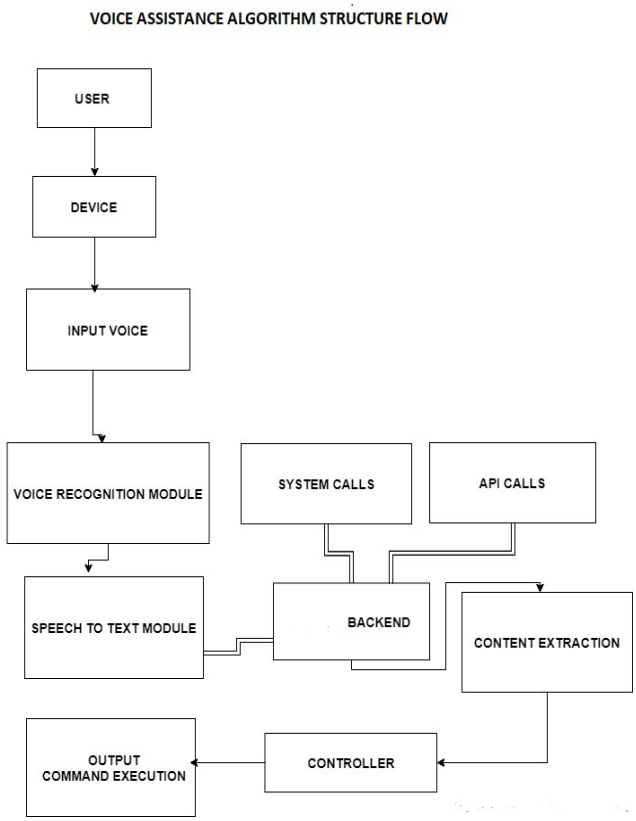


Fig. 4. Algorithm for Voice Assistant.

B. Voice Assistant Algorithm

1. Initialize the device by calling its name.
2. Speech Recognition - The system will use Google's online speech recognition system for converting recorded speech input to text.
3. Backend - The backend will receive the text output from speech recognition module and check if the command is an API call, system call or content extraction.
4. Command execution The command requested by the user is executed and the output is the send to the user.

VI. DISCUSSION

The purpose of this paper is to suggest changes in the traditional way (how a restaurant is managed) so that the negative impact of COVID-19 in the food industry can be mitigated.

One of the changes is to have a food ordering system in a restaurant that will be scalable and can act as a great coordinator helping staff get align for better customer service.

This project also aims to reduce point of contact during COVID-19 and automate the ordering and billing process, simplifying the work. Reducing paper waste, making records digital for better reliability and management. Allow restaurant to run with minimum workforce. Reduce restaurant-running cost and allow the restaurant to run while smoothly while following COVID norms.

A. Future Work

- The system can be made compatible with different forms of payment like debit cards, UPI, digital wallets, etc. The UI can be improved.
- We can also add a feature to see live order status and provide deals to influence customer regularity.
- Apart from the voice assistance, we add a module of image processing with the help of open CV to

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:38:49 UTC from IEEE Xplore. Restrictions apply.

predict the age and gender of the customer and suggest better options in the menu. Making the interface more interactive and smarter. Open CV module can also be used to check for facemask and suggest customer to wear facemask after they are done eating.

The accuracy of voice recognition can be improved.

B. Limitations

- The initial Setup and the learning curve can pose a bit of a problem.
- The suggested system is connected via Wifi router. So the security of the Wifi can pose a challenge. The wifi router can also act as a point of failure that could cripple the system.
- The voice recoginision could produce incaaurate results sometimes.

VII. CONCLUSION

The suggested system and changes will help the restaurant work at its full capacity with minimum workforce following all the COVID-19 guidelines. The new system will reduce the burden on the workforce by ensuring only necessary interaction takes place and proper social distancing is maintained and COVID-19 guidelines are followed.

Merits

- Customers will not have to line up to place orders
- Social media marketing will increase popularity/presence while keeping the advertisments cost down.
- Profitibality of the restaurant will increase

Demerits

- Staff will have to be initally trained to use the new system.
- System maintainence will require someone with techinical knowladge.
- Security of the new system can pose a challenge.

ACKNOWLEDGMENT

We would like to acknowledge the contributions of the following people, without whose help and guidance this report would not have been completed. We acknowledge our mentor, Mr. Ram Murti Rawat, Assistant Professor, Depart of Computer Science & Engineering, Delhi Technological University, with respect and gratitude, whose expertise, guidance, support, encouragement, and enthusiasm has made this report possible. Their feedback vastly improved the quality of this report and provided an enthralling experience. We are indeed proud and fortunate to be supported by him. Although it is not possible to name individually, We shall ever remain indebted to all the faculty members of Delhi Technological University, for their persistent support and cooperation extended during this work.

This acknowledgment will remain incomplete if I fail to express our deep sense of obligation to my parents and God for their consistent blessings and encouragement.

REFERENCES

[1] "COVID-19 Era Serves Up Big Changes For U.S. Restaurants | Morgan Stanley".2021 MorganStanley. Available Online: https://www.morganstanley.com/ideas/coronavirus-restaurant-trends.

[2] "Food Security And COVID-19". 2021. World Bank. Available Online : https://www.worldbank.org/en/topic/agriculture/brief/food-security-and-covid-19.

[3] Bhadane, Satish, Shweta Vyapari, Aditi Potdar, Ragini Salve, and Sukanya Jorvekar. 2017. "RESTURANT AUTOMATION SYSTEM USING ANDROID". International Journal Of Recent Trends In Engineering And Research 3 (4): 604-607. doi:10.23883/ijrter.2017.3200.lwc0l.

[4] Nakat, Zeina, and Christelle Bou-Mitri. 2021. "COVID-19 And The Food Industry: Readiness Assessment". Food Control 121: 107661. doi:10.1016/j.foodcont.2020.107661

[5] E.Abel, Ejde, and Ekabua Obeten. 2015. "Restaurant Customer Self-Ordering System: A Solution To Reduce Customer/Guest Waiting Time At The Point Of Sale". International Journal Of Computer Applications 111 (11): 19-22. doi:10.5120/19583-1332.

[6] Graham Byford, Roger. 2013. "VOICE ASSISTANT SYSTEM". The Journal Of The Acoustical Society Of America 133 (4): 2520. doi:10.1121/1.4800165

[7] Maxime Lemerle. July 2020. " CALM BEFORE THE STORM: COVID-19 AND THE BUSINESS INSOLVENCY TIME BOMB". Euler Hermes, Allianz Research

[8] Kumar, N.S., 2019. Implementation of artificial intelligence in imparting education and evaluating student performance. Journal of Artificial Intelligence, 1(01), pp.1-9.

[9] T. Kim, "Short Research on Voice Control System Based on Artificial Intelligence Assistant," 2020 International Conference on Electronics, Information, and Communication (ICEIC), Barcelona, Spain, 2020, pp. 1-2, doi: 10.1109/ICEIC49074.2020.9051160.

[10] Ranganathan, G., 2020. Real time anomaly detection techniques using pyspark frame work. Journal of Artificial Intelligence, 2(01), pp.20-30.

[11] Szmigiera, M., 2021. "Impact of the coronavirus pandemic on the global economy - Statistics & Facts". Statista. Available Online : https://www.statista.com/topics/6139/covid-19-impact-on-the-global-economy/

[12] Karjaluoto, Heikki, Aijaz A. Shaikh, Matti Leppäniemi, and Roope Luomala. "Examining Consumers' Usage Intention of Contactless Payment Systems." International Journal of Bank Marketing 38, no. 2 (November 18, 2019): 332–351. doi:10.1108/ijbm-04-2019-0155

[13] Rizou, M, Galanakis, I M, Aldawoud, T M S, et al. (2020). Safety of foods, food supply chain and environment within the COVID-19 pandemic. Trends in Food Science & Technology, 102:293–299.

[14] FAO (Food and Agriculture Organization of the United Nations). (2020j). Policy responses to keep input markets flowing in times of COVID-19. Available Online : http://www.fao.org/3/ca8979en/CA8979EN.pdf.

[15] Bakalis, S, Valdramidis, V P, Argyropoulos, et al. (2020). Perspectives from CO+RE: how COVID-19 changed our food systems and food security paradigms. Current Research in Food Science, 3:166–172

[16] Hobbs, J E. (2020). Food supply chains during the COVID-19 pandemic. Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie, 68:171–176.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:38:49 UTC from IEEE Xplore. Restrictions apply.

JOURNAL OF MECHANICAL ENGINEERING AND SCIENCES (JMES)
ISSN: 2289-4659 e-ISSN: 2231-8380
VOL. 15, ISSUE 2, 8057 – 8071
DOI: https://doi.org/10.15282/jmes.15.2.2021.10.0635

ORIGINAL ARTICLE

An FEA based study of thermal behaviour of ultrasonically welded phosphor bronze sheets

Bharat Sanga[1], Reeta Wattal[2] and D.S. Nagesh[3]

[1] Department of Mechanical Engineering, Guru Nanak Dev Institute of Technology, Delhi 110089, India
Phone: +919013310266
[2,3] Department of Mechanical Engineering, Delhi Technological University, Delhi 110042, India

ABSTRACT – The ultrasonic joining of phosphor bronze sheets is analyzed using a 3-D finite element model for the study and prediction of the thermal profiles at the weld interface. The heat fluxes are calculated and assigned as boundary conditions during the thermal simulation. The forecast of temperature is done under various welding conditions. The maximum temperature obtained by transient simulation at the weld interface is 366.74℃. The continuous reduction in the temperature is observed towards the extremes of the weld metal. The sonotrode and the anvil achieve a lower temperature in comparison to the weld interface. The effect of clamping force and bonding ratio on the interface temperature is observed as positive. The model is validated with an error of 1.576% between the observed and predicted temperature results and a correlation co-efficient 0.96 is established between the simulated temperature results and the weld strength. Sufficiently strong joints were obtained at the optimum welding conditions with 74% joint efficiency. It is evident that the interface temperature has a strong linear relationship with joint strength and is a major deciding factor for achieving strong joints.

ARTICLE HISTORY
Received: 24th July 2020
Revised: 11th Oct 2020
Accepted: 24th Jan 2021

KEYWORDS
Ultrasonic welding;
finite element analysis;
simulation;
SimScale;
heat flux;
temperature profile;
plastic deformation;
friction;
thermocouple

INTRODUCTION

The current industrial need of producing innovative and sustainable products demands for robust and compact design. The need can be accomplished by the use of light-weight materials fulfilling specific product design requirements. Typically, aluminum and copper based alloys are extensively chosen for manufacturing parts of Battery Electric Vehicles (BEV) such as electrical contacts, terminals, current collectors, and bus bars owing to their excellent strength and spectacular electrical and thermal properties [1] . Some commonly used joining techniques such as resistance welding, micro-TIG welding and laser welding pose electrical, thermal, metallurgical, and mechanical challenges to the joint characteristics [2]. Incessant efforts are being made to evolve better techniques for thin sheet joining. Conventional joining methods are modified to suit the needs of heterogeneous tiny joints. Resistance micro-welding, laser droplet joining, duo-thermal soldering process, and precision miniaturized droplet deposition technique are some of the newly evolved techniques [3–6]. These techniques still need standardization due to diverse material characteristics. Ultrasonic welding (USW) is an emerging technique for manufacturing these parts owing to its superfast joining time, excellent weld quality, and capability to join dissimilar materials with the least power consumption. It is a friction oriented joining method that utilizes acoustic energy to plastically deform the metal and bringing the nascent metal surfaces in close contact so as to form a strong metallic bond [7]. Fairly seen simple from outside, the actual mechanism of joint formation in USW is quite complex [8]. In fact, an accurate and well-established theory related to the joint mechanism of USW is yet to be developed [9]. There is a need to regulate the operating procedure to make this technique a commonly utilized popular joining method.

Process optimization and microstructural evaluation of USW are some of the fields explored by a number of researchers [10–13] for a variety of metal combinations. Elangovan et al. [10] joined dissimilar 0.2 and 0.3 mm thick sheets of Cu and brass using USW. Welding parameters were optimized to maximize weld strength by interfacing the Response Surface Method (RSM) with the Genetic Algorithm (GA). Similarly, thin sheets of aluminum and steel were successfully joined and parameters were optimized using Artificial Neural Network (ANN) in association with GA by Zhao Dewang et al. [11]. During their observation, weld time was found to be the most important parameter, however, vibration amplitude and clamping force did not affect significantly. A dissimilar weld between aluminum and copper with aluminum alloy particle interlayer was created and the effect of weld energy on weld strength was investigated and substantiated with the microstructural evolution by Ni Zenglei et al. [12]. The weld strength increased with increasing energy and no traces of IMCs were observed. The particle interlayer also affected the weld strength to a great extent. Copper wire joints find numerous applications in electrical and electronics products. Mohan Raj et al. [13] investigated the joining of copper wire with the copper substrate. Their work involved parametric optimization, microstructural evaluation, and numerical modelling. The FEM analysis revealed that thermal stresses developed at the interface. FEM based models are simulated for estimating mechanical and thermal impacts of operating conditions of ultrasonically

*CORRESPONDING AUTHOR | Bharat Sanga | ✉ bharatsanga_2k14@dtu.ac.in

© The Authors 2021. Published by Penerbit UMP. This is an open access article under the CC BY license.

welded joints [14-21]. Zhong et al. [14] performed an experimental study and modelling of the ultrasonic consolidation process of dissimilar Ti and Al alloys. The highest temperature was obtained at the common surfaces of foil-sonotrode and foil-substrate. They also observed that the oscillator amplitude had more effect on the interfacial temperature than the consolidation pressure. Besides metals, modelling of thermo-plastic was performed using ANSYS software and the results were authenticated with the experimentally obtained data [15]. A detailed study of different similar and dissimilar metal ultrasonic joints was done by DeVries [16]. The effect of significant process parameters, microstructural transformation, and evaluation of HAZ in the weld zone and simulation of 3-D CAD model were the primary objectives of his work. The heat developed at the weld interface was estimated and the model was utilized to predict the interfacial temperature. Similar work was done by Elangovan et al. [17, 18]. The finite element based model was presented to simulate the actual working conditions of the ultrasonic welding for predicting the temperature distribution at the weld interface as well as the stress distribution in the sonotrode and the joint. The simulated results were found in good concurrence with those acquired experimentally. Chen and Zhang [19] constructed and simulated the 3-D model of USW of dissimilar automotive alloy and found that the temperature distribution was not uniform at the knurled surface of the sonotrode tip. The amount of heat produced due to plastic deformation was almost one-fourth of the total heat produced at the interface and the highest temperature was located at the weld interface. Aluminum alloy AA 6011 was joined separately by magnesium alloy AZ31 and low carbon steel DC04 by Jedrasiak et al.[20]usingUSW. The heat input during welding was evaluated using thermocouple data and some other evidence. The thermal histories were utilized to predict the thickness of the intermetallic compound. A 3-D finite element model of the thermal field was applied for temperature prediction at the weld interface. The same author also simulated the heat generation in USW using finite element analysis [21]. A single cycle oscillation deformation model was employed for calculating the heat generation rate sporadically. A continuous thermal model was capable to predict the temperature field.

It is clear from the review of the literature that modelling of the thermal field plays an important role in the prediction of the temperature profile at the weld interface. The response characteristics of the USW process can be forecast, and the corresponding data can be utilized optimally by the industry. Keeping these implications in mind, the current work is focused on the development and simulation of a 3-D finite element model of thin sheets of phosphor bronze joined by USW. The CAD model was employed to have a close inspection into the thermal phenomenon leading to the joining process and consequent assessment of its impact. This work involves the following objectives: (i) measurement of interface temperature and the strength of the weld joint (ii) calculation of heat generation at the weld interface (iii) development of a CAD model and its analysis using FEA technique to predict the temperature at different zones of the interface and (iv) establish a correlation between the interface temperature and the strength of the joint.

METHODS AND MATERIALS

Equipment and Material Layout

In the present study, the weld coupons of phosphor bronze sheets (UNS C51100) having size 100 mm x 25 mm x 0.36 mm have been prepared with a 25 mm overlap. A lateral drive Telesonic® ultrasonic welding machine (20 kHz, 3 kW) with knurled horn tip of size 13 mm x 10 mm was employed for performing the experimental work. The experiments and simulation are performed considering three parameters, weld time (W_t) clamping force (F_c), and vibration amplitude of the sonotrode (V_a). The selection of parameters is based on the literature survey and the preliminary experiments conducted to evaluate the range of parameters. The value of each parameter was varied at three levels as shown in Table 1.

Table 1. Welding parameters and their levels

Factor	Unit	Symbol	Level 1	Level 2	Level 3
Weld Time	Sec	W_t	0.50	0.85	1.20
Clamping Force	N	F_c	1100	1650	2200
Vibration Amplitude	µm	V_a	28.0	31.5	35.0

The ultimate tensile-shear load of the weld joint and the temperature at the interface were measured to evaluate the performance of the ultrasonic welded joint. The experiments were executed as per the Box-Behnken design owing to its advantage of requiring a lesser number of experiments. Each weld parameter was selected at 3 levels for conducting 15 randomized experiments each with 6 replicates. Before making the joints, the specimens were properly swabbed with acetone and dried to remove the oils and grease from the contacting surfaces to ensure close contact between them. The tensile-shear load of the weld coupons was tested on Tinius Olsen (Model: H50KS) computerized universal testing machine (UTM) with constant crosshead speed at 1 mm/min. The schematic diagram and the actually welded coupons are shown in Figure 1.

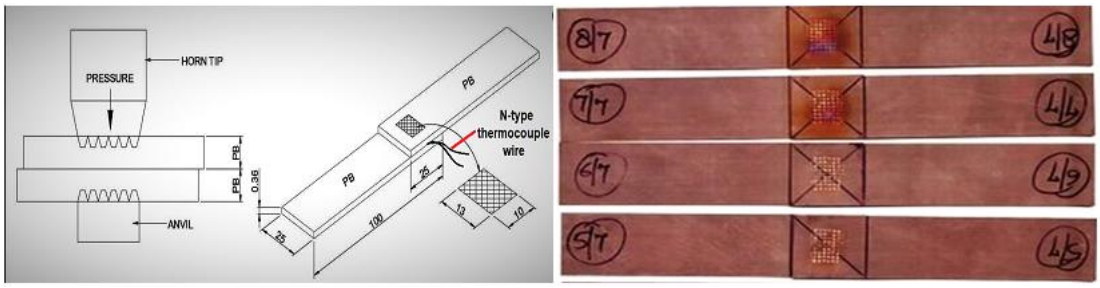


Figure 1. Schematic of weld coupon with N-type thermocouple wire attached at the middle of the overlapped area and experimental weld specimens

The real-time interface temperature was measured using a Data Acquisition System (DAQ) during the welding. Data Translation® make DT9828, 8 channel isolated thermocouple DAQ module having a sampling rate of 600 Hz and accuracy of 0.09℃ was employed for the processing of temperature data. The temperature measuring sensor used for this work was 0.20 mm diameter sacrificial N-type (Nicrosil-Nisil) thermocouple wire which has a working range between -270℃ and +1300℃. It was placed at the centre of the overlapped portions of the weld coupon touching the interface, considering it as the hottest point during the welding. The N-type thermocouple has a very high sensitivity index (39 µV / ℃ at 900℃) and thermo-electrical stability as compared to other thermocouples. Further, the DAQ module was interfaced with the QuickDAQ software to process the real-time interface temperature value. The QuickDAQ software has the functionality of converting the electrical signals generated by the thermocouple to the temperature values and showing them in the graphical form. The whole system of temperature measurement and processing can be shown through a block diagram as shown in Figure 2.

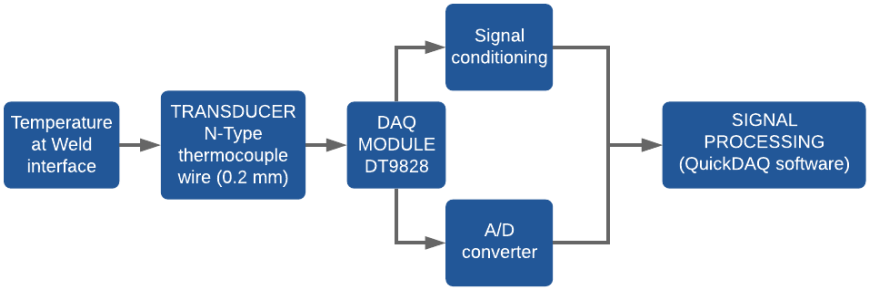


Figure 2. Block diagram of DAQ system

THERMAL ANALYSIS OF ULTRASONIC WELDING

The mechanism of joint formation in USW makes it a complicated joining process due to the application of remarkably low-amplitude and high-frequency vibrations on a small region for a very short period. The heat produced at the joining surfaces is not ample to melt the metal but again is sufficient enough to produce a strong joint in the cold condition. The contaminants and the surfacial asperities initially prevent the weld metals to make intimate contact. But at the same time, the periodic shear stresses developed due to the vibration of the sonotrode and the pressure exerted on it produce sufficient instantaneous heat. It eventually softens the metal, disperses the contaminants and thus, a sound metallic joint is formed. The factors responsible for the rise of a significant amount of heat delivered at the weld zone during ultrasonic welding are material deformation, friction, and vibration [22]. Initially, the effect of friction is observed between the sonotrode tip and the upper sheet for negligible time, hence, this may be considered insignificant [16].

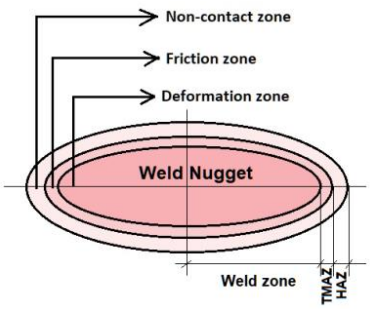


Figure 3. Weld nugget demarcation into different zones

In the meantime, as the welding process proceeds, the heat due to deformation and friction between the sheets rises significantly and the formation of small micro-welds begins. Very soon, these micro-welds expand and propagate on a larger surface. The form of the micro-welds then converts into macro-welds or the weld nuggets. Figure 3 shows the different zones of the affected area.

The zone where deformation of the weld metal takes place and formation of the micro-welds starts is known as the 'deformation zone'. The size of this zone increases continuously and may reach up to the size of the sonotrode tip depending upon the penetration of the sonotrode teeth. The effect of heat and stress is observed in this zone. This zone is surrounded by the 'friction zone' whose size is slightly more than the deformation zone due to the vibrating sonotrode. Although the weld metal gets deformed in this zone, the formation of micro-welds does not take place due to insufficient heat generation. This zone is also known as the thermo-mechanically affected zone (TMAZ). The deformation of metal and the friction between the faying surfaces are the primary factors of heat generation in the 'deformation zone' and the 'friction zone' respectively. Lastly, the heat-affected zone (HAZ) is found adjacent to TMAZ which is characterized by the heat only [23]. This work involves the well-approximated estimation of the heat generation due to the deformation of the metal and friction between the mating surfaces. Based on these estimations, a thermal model has been prepared and utilized to calculate the size of the weld zone, TMAZ, and HAZ.

Measurement of Weld Area

The estimation of the actual weld area is essential for the calculation of heat flux at the interface. The weld joint formed in the case of USW contains a number of small macro-welds or nuggets. The net weld area is the sum of all the nuggets. It was observed that the weld coupons stick on to a relatively smaller portion of the interface that makes it difficult to estimate the actual weld area. The actual weld area is the result of different combinations of parameters. In the case of knurled sonotrode tip and thin weld metal sheets, the depth by which the teeth on sonotrode tip get penetrated onto the upper sheet vary with parameter combination. Secondly, unlike resistance spot welding, the weld area is the sum of the areas of a number of micro-welds which appear on the separated surfaces as macro-welds or weld spots. The metal under a tooth gets deformed and spreads in a larger surface, but as observed, it does not join strongly with its counter-surface surrounding the teeth. Whereas, the interface below a tooth gets attached so strongly that while detaching it peels off the surface of another sheet.

Two different methods have been employed to enhance the accuracy of the work due to the irregular shape and size of the weld spots. In the first method, the cross-sectional area of one sonotrode tooth was calculated by enlarging the tooth impressions on the interface using a profile projector and multiplied by the number of weld spots present at the weld interface. In the second method, the bond impressions on the magnified interface were summed up using ImageJ software (Version 1.52 a). Both sides of the interface of each weld coupon were taken into consideration to improve the precision of the results. This way, six areas were measured at each run of the design matrix. The mean value of the two areas was taken as the final weld area.

Heat Flux Calculation

The basic model for the calculation of heat flux is based upon the fact that the joint is formed as a result of heat produced due to (i) the shear deformation of the weld metal below the sonotrode tip and (ii) the friction between the contacting surfaces of the weld metal. The shear deformation of the weld metal is subjected to (i) the normal compressive stresses due to clamping force and (ii) the shear stresses due to ultrasonic vibrations. The combined effect of these two stresses brings the material under plastic deformation resulting in the generation of excessive heat which raises the temperature of the material instantly. As the deformation takes place under one normal and one shear stress, the failure of the material may be considered to follow Tresca's failure criterion which states that failure occurs when the maximum shear stress reaches its critical value that is the yield stress in a normal tensile test. Using this concept, the value of shear stress which is responsible for the deformation may be calculated as:

$$f_{max} = \frac{\sigma_1 - \sigma_2}{2} = \sqrt{\left(\frac{\sigma_N}{2}\right)^2 + (f_s)^2} \quad (1)$$

Where, f_{max} is the maximum shear stress, σ_1 and σ_2 are principal stresses, σ_N is the normal stress due to pressure and f_s is the shear stress responsible for deformation. So, according to Tresca's condition, if 'Y' is the yield stress of the material during a normal tensile test, Eq. (1) can be formulated as:

$$\frac{Y}{2} = \sqrt{\left(\frac{\sigma_N}{2}\right)^2 + (f_s)^2} \quad (2)$$

So, value of shear stress,

$$f_s = \sqrt{\left(\frac{Y}{2}\right)^2 - \left(\frac{\sigma_N}{2}\right)^2} \quad (3)$$

Or

$$f_s = \sqrt{\left(\frac{Y_t}{2}\right)^2 - \left(\frac{F_c / A_d}{2}\right)^2} \tag{4}$$

Where, Y_t is the temperature-dependent yield stress, F_c is the clamping force and A_d is the area of the deformation zone. The power dissipated (P) in the deformation zone produces heat flux at the interface which is a function of the average sonotrode velocity (V_{avg}) and the weld force (F_w). Therefore, the heat flux due to deformation (H_d) in W/m^2 can be shown as Eq. (5):

$$H_d = \frac{P}{A_d} = \frac{F_w \times V_{avg}}{A_d} \tag{5}$$

Ideally, for thin sheets and flat sonotrode tip, the area of deformation can be considered equal to the sonotrode area (A_s) since the entire area beneath sonotrode tip gets deformed. Now, Eq. (5) can be written as:

$$H_d = \frac{f_s \times A_w \times V_{avg}}{A_s}$$
$$H_d = (A_w / A_s) \times V_{avg} \times f_s \tag{6}$$

Substituting the shear stress, f_s in Eq. (6) from Eq. (4), we can write Eq. (7) as

$$H_d = (A_w / A_s) \times V_{avg} \times \sqrt{\left(\frac{Y_t}{2}\right)^2 - \left(\frac{F_c / A_s}{2}\right)^2} \tag{7}$$

$$H_d = \left(\frac{\beta}{2}\right) \times V_{avg} \times \sqrt{(Y_t)^2 - (F_c / A_s)^2} \tag{8}$$

The average velocity of the sonotrode in Eq. (8) can be calculated by $V_{avg} = 4 \times V_a \times V_f$ where V_a and V_f are the amplitude and frequency of the ultrasonic vibrations. 'β' is the bonding ratio which signifies the extent of the area under the sonotrode tip that is transformed into weld zone. Its value is indicative of the effectiveness of the process which itself is the resultant of many factors and one of the most important factors is the heat generated at the interface. The ideal value of 'β' should be 1 when the whole of the area under sonotrode tip is engrossed with the micro-welds but due to the presence of oxide, asperities, and other contaminants, some portion remains unbounded.

The temperature-dependent yield strength (Y_t) of phosphor bronze is assumed to equal to its yield strength in normal conditions. Some alloys such as phosphor bronze do not lose an appreciable amount of their strength at a higher temperature until 350℃ [24] and it was observed during experimentation that the maximum average interface temperature where deformation takes place did not exceed 354.7℃. The calculation of heat flux due to deformation of weld metal according to Eq. (8) has been done for weld coupon having the lowest bonding ratio (β) of 0.26, clamping force of 1100 N, weld time of 0.85 sec, sonotrode amplitude of 28 µm and is shown below:

$$H_d = \left(\frac{0.26}{2}\right) \times 2.24 \times \sqrt{(295 \times 10^6)^2 - (1100/(130 \times 10^{-6}))^2} = 85.87 \times 10^6 \text{ W/m}^2$$

The heat flux due to friction (H_f) between the mating surfaces of the workpiece depends upon the dynamic coefficient of friction between the surfaces (μ_k), the clamping force (F_c), and the average sonotrode velocity (V_{avg}). The value of μ_k is taken as 0.85 for phosphor bronze sheets. Heat flux due to friction may be calculated using Eq. (9).

$$H_f = \frac{\mu_k \times F_c \times V_{avg}}{A_s} \tag{9}$$

Therefore, $$H_f = \frac{0.85 \times 1100 \times 2.24}{130 \times 10^{-6}} = 16.11 \times 10^6 \text{ W/m}^2$$

The total heat flux H is the sum of heat flux due to deformation (H_d) and heat flux due to friction (H_f). Therefore, H = 85.87 x 10^6 W/m^2 + 16.11 x 10^6 W/m^2 = 101.98 x 10^6 W/m^2. A similar calculation was performed for calculating the heat fluxes at each run and the heat flux value has been imposed as one of the boundary conditions during simulation of the CAD model of the USW of phosphor bronze sheets.

Finite Element Analysis (FEA)

The CAD model of the ultrasonic welded joint between phosphor bronze sheets was prepared and simulated for thermal analysis using SimScale Workbench 2.0 software (SimScale GmBH). SimScale Workbench is an open-source cloud-based CAE computing environment providing up to 96 core computing capability for simulation [25]. The geometry, mesh, and the working boundary conditions of the FE model are shown in Figure 4. The proposed model can be used for simulation of USW of similar as well as dissimilar materials. The CAD model, which is actually a replica of USW process, is comprised of three parts connected closely with each other - the weld metals, the anvil and the sonotrode. Accordingly, the model was associated with three materials: the weld material phosphor bronze (UNS C51100), sonotrode material Ti alloy, and anvil material tool steel. The complete assembly of sonotrode was not included in the model, as the thermal conductivity of the sonotrode material is very low as compared to the weld metal and the processing time in USW is very small. Hence, only the sonotrode tip is included in the analysis. The meshing of the model was kept to the finest with tetrahedral mesh element graded from coarse to fine, with the finest mesh at the central part. The current model used 577856 elements with 152203 nodes and 4164 edges as shown in Figure 4. This meshing scheme was adopted to improve the computational accuracy of the analysis and to reduce the hourglass effect [26].

Along with that, the physical, mechanical and thermal properties of the materials such as melting point, Young's modulus, yield strength, Poisson's ratio, thermal conductivity, etc. are included in the model [27] and are shown in Table 2. All the components of the FE model are presumed equivalent to the actual physically assembled USW process in their orientation and physical attribute. The anvil, a stationary part of the ultrasonic metal welding machine is made fixed in the FE model also by restricting all of its degrees of freedom. The clamping pressure plays an important role in the formation of the weld joint in USW. So, it becomes imperative to give due consideration to its effect on different contact areas in the FE model. The portions of the weld coupons which are away from the thin weld zone and are surrounded by air may be considered as insulated due to the low convective coefficient of air and short weld time. Similarly, it was practically observed that the portions of the overlapped weld coupons away from the weld zone get slightly separated due to less impact of clamping pressure, and the air gets entrapped between the sheets. Hence, they are also treated as insulated boundaries. The contacting parts must have a large thermal contact conductance being highly conductive soft metals with clean and flat faying surfaces. Pressure as well as vibratory motion also contribute to enhance its value [28]. So, the value of thermal contact conductance was taken as 10^4 $Wm^{-1}°C^{-1}$ for the weld interface and the interface between the sonotrode tip and the upper weld coupon [20].

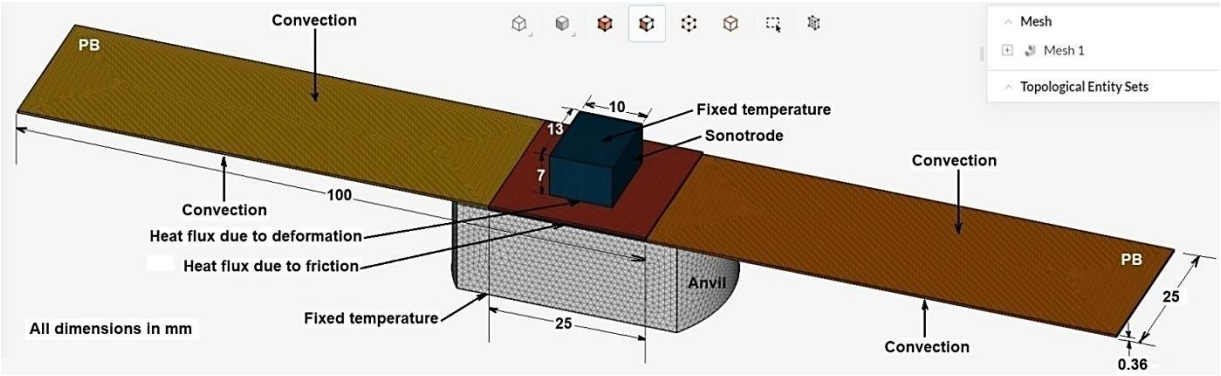


Figure 4. Boundary conditions imposed in the three dimensional FE model of USW

Table 2. Material properties

Property	Unit	Weld material – UNS C51100	Sonotrode Material – Ti alloy (Ti-6Al–4V)	Anvil Material – Tool Steel
Density	Kg/m^3	8860	4430	7670
Young's modulus	GPa	110	113.8	205
Yield Strength	MPa	295 (H01 hard)	880	-
Poisson's ratio	-	0.34	0.34	0.30
Thermal conductivity	W/m.K	84	6.7	20
Specific heat	J/kg.K	380	526.3	460

The heat flux values calculated according to the Eqs. (5) and (6) were applied to the interface only rather than to the entire assembly. It was done according to the work of Jedrasiak [29] who established that in the case of thin metallic deforming sheets, the heat fluxes applied separately to the interface and the bulk, do not affect much to the temperature distribution at the interface. The heat flux produced due to deformation of weld metal was applied uniformly at the

interface where the sonotrode tip touches the upper weld coupon whereas the heat flux due to friction was applied at the contacting weld interface. The upper surface of the sonotrode tip and the bottom surface of the anvil, both were kept at a fixed temperature of 30℃ assuming there will be no heat flux across these surfaces in this short weld time. The outer surfaces of the model which were exposed to the air and not in contact with the sonotrode or the anvil were assigned with natural convection having the convective heat transfer coefficient of 5 $Wm^{-2}℃^{-1}$[17]. The ambient temperature was taken as 30℃ and the whole model was set at this initial temperature. The simulation was executed in a transient condition taking a time step of 0.1 sec.

RESULTS AND DISCUSSION

The thermal simulation study of the CAD model of USW of phosphor bronze sheets was performed, and the prediction results are tabulated along with the experimentally observed values of the interface temperature in Table 3. All precautions were taken during temperature measurement, despite that some results were found inconsistent due to extremely fast processing time, the possibility of thermocouple deformation and vibration at the point of measurement. The effect of tremendous vibrations on the temperature profile was visible in the enlarged view, where the temperature values are fluctuating in the range of 10-15 ℃ for a period of about 0.011 seconds as shown in Figure 5. Therefore, the measurements were repeated until the reliable results were obtained and only those temperature profiles were accepted which have similar heating and cooling histories without any abnormal fall or quick rise.

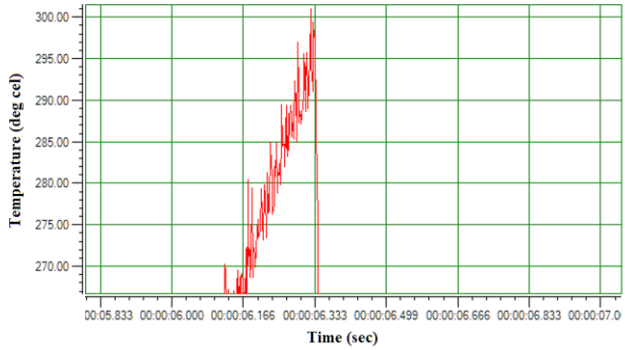


Figure 5. Fluctuation of temperature due to vibration during ultrasonic welding

Table 3. Experimental and simulation temperature data

Run No.	Clamping Force (N)	Sonotrode Amplitude (µm)	Weld Time (sec)	Weld Area (mm^2)	Bonding ratio (β)	Simulation Temperature data (℃)	Experimental Temperature data (℃)			Average Temperature (℃)	Tensile-shear load* (N)
							Trial 1	Trial 2	Trial 3		
1	1650	28.0	1.20	65.40	0.42	321.3	329.7	292.6	319.4	313.9	1580.00
2	1650	31.5	0.85	51.30	0.50	315.0	301.0	327.1	315.1	314.4	1781.67
3	2200	28.0	0.85	71.58	0.39	294.0	293.9	299.1	300.0	297.7	1226.67
4	1100	31.5	0.50	41.25	0.43	244.0	239.3	258.3	223.6	240.4	603.00
5	1100	28.0	0.85	33.78	0.26	250.5	255.9	263.8	243.7	254.5	703.67
6	1650	35.0	1.20	56.40	0.57	351.7	345.6	335.2	340.6	340.5	2306.67
7	1650	35.0	0.50	55.00	0.38	287.2	295.3	308.2	295.4	299.6	1246.67
8	1100	35.0	0.85	74.33	0.45	276.9	293.5	275.2	270.2	279.6	738.33
9	1650	28.0	0.50	59.08	0.32	281.0	290.0	294.9	283.5	289.5	1216.00
10	2200	35.0	0.85	76.48	0.65	366.7	370.7	332.6	361.0	354.7	2353.33

Table 3. Experimental and simulation temperature data (cont.)

Run No.	Clamping Force (N)	Sonotrode Amplitude (μm)	Weld Time (sec)	Weld Area (mm^2)	Bonding ratio (β)	Simulation Temperature data (℃)	Experimental Temperature data (℃)			Average Temperature (℃)	Tensile-shear load* (N)
							Trial 1	Trial 2	Trial 3		
11	2200	31.5	1.20	84.88	0.59	346.4	370.5	330.8	346.8	349.4	2306.67
12	1650	31.5	0.85	49.88	0.55	311.3	316.9	327.8	313.7	319.5	1853.33
13	2200	31.5	0.50	63.40	0.39	311.9	297.5	296.3	311.0	301.6	1396.67
14	1650	31.5	0.85	50.85	0.49	305.5	318.9	315.7	304.1	312.9	1666.67
15	1100	31.5	1.20	45.05	0.35	295.7	293.1	280.2	293.4	288.9	1196.67

* The mean of three experimental trials.

The first simulation trial was performed at run number 5 which has the smallest value of bonding ratio (β = 0.26) with clamping force 1100 N, sonotrode amplitude of 28.0 μm and, weld time 0.85 sec. The values of two process parameters, the clamping force and sonotrode amplitude were assigned at a low level in that run. The heat flux values H_d and H_f were 85.87 x 10^6 W/m^2 and 16.11 x 10^6 W/m^2 respectively. The simulation trial performed with those inputs returned the temperature of 250.5℃. The value of the peak average interface temperature obtained by the thermocouple at the same parameter settings was 254.5℃. With those parameter settings, the tensile-shear load was obtained as 703.67 N. The values of interface temperature and load are towards the lowest side at the low level of parameters and bonding ratio. The corresponding results from real-time observation and simulation are shown in Figure 6. The temperature distribution is shown along the longitudinal section (x-axis) of the assembly.

The second simulation trial was performed at run number 14 which has the moderate value of bonding ratio (β = 0.49) with clamping force 1650 N, sonotrode amplitude of 31.5 μm, and weld time 0.85 sec. The values of all the process parameters were assigned at the middle level in that run. The heat flux values H_d and H_f were 181.96 x 10^6 W/m^2 and 27.19 x 10^6 W/m^2 respectively. The simulation trial performed at these inputs returned the temperature of 305.5℃. The value of the peak average interface temperature obtained by the thermocouple at the same parameter settings was 312.9℃. The tensile-shear load obtained at this parameter setting was 1666.67 N. The values of interface temperature and load value are higher than the previous simulation run. The corresponding results from real-time observation and simulation are shown in Figure 7. The temperature distribution is shown along the longitudinal section of the assembly.

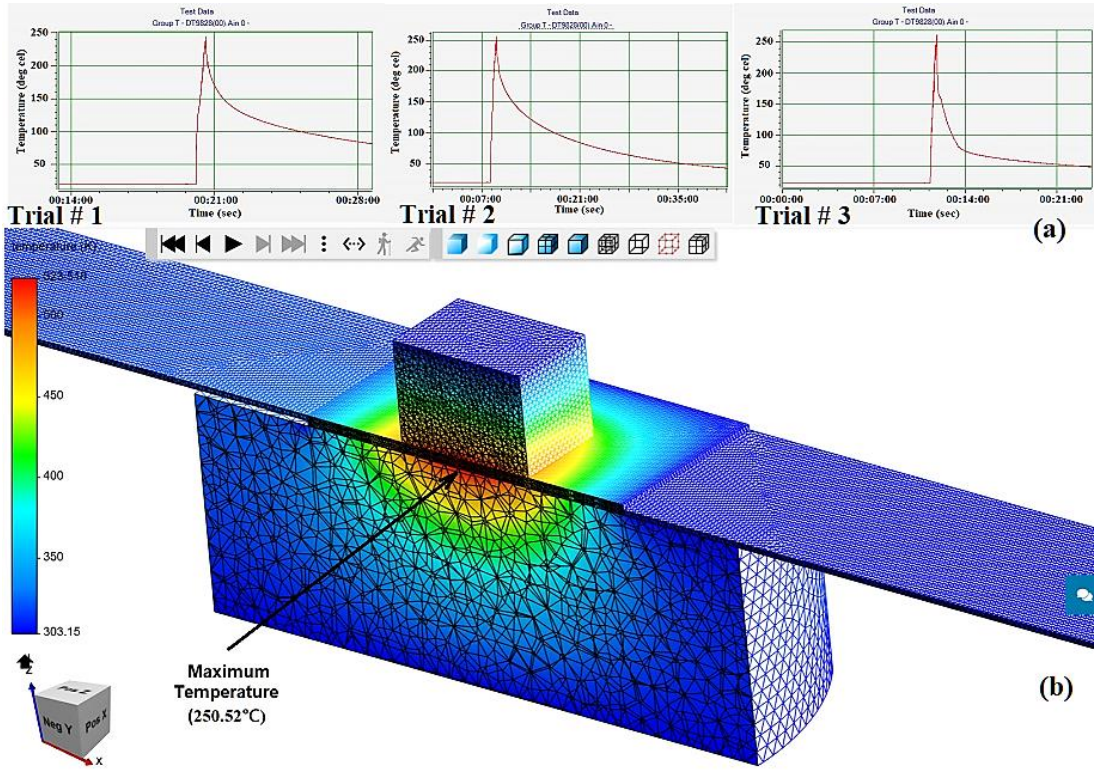


Figure 6. (a) Temperature profile of experimental results and (b) simulation result of run 5

The third simulation trial was performed at run number 10 which has the highest value of bonding ratio ($\beta = 0.65$) with clamping force 2200 N, sonotrode amplitude of 35.0 µm, and weld time 0.85 sec. The values of two process parameters, the clamping force and sonotrode amplitude were also assigned at a high level in that run. The heat flux values H_d and H_f were 268.01 x 10^6 W/m^2 and 40.28 x 10^6 W/m^2 respectively. The simulation trial performed with those inputs returned the temperature of 366.7℃. The value of the peak average interface temperature obtained by the thermocouple at the same parameter settings was 354.7℃. The tensile-shear load obtained at that parameter setting was 2353.33 N. The values of interface temperature and load are the highest at the high level of parameters and bonding ratio. The corresponding results from real-time observation and simulation are shown in Figure 8. The temperature distribution is shown along the longitudinal section of the assembly.

Further, it was observed from the simulated results that the interface temperature increased with the increased clamping force while keeping the sonotrode amplitude and weld time constant. Run 5 and run 3 were simulated at a low level, 28.0 µm of sonotrode amplitude. Keeping the weld time fixed at 0.85 sec, the simulated results of interface temperature and the value of tensile-shear load increased from 250.5℃ to 294.0℃ and 703.67 N to 1226.67 N respectively, when the clamping force was increased from 1100 N to 2200 N. Run 4 and run 13 were simulated at the middle level of sonotrode amplitude at 31.5 µm and fixed value of weld time at 0.5 sec. The interface temperature and the tensile-shear load increased from 244℃ to 311.9℃ and from 603 N to 1396.67 N respectively, when the clamping force was increased from 1100 N to 2200 N. Similarly, high level 35.0 µm of sonotrode amplitude and fixed value of weld time at 0.85 sec. were used for simulating run 8 and run 10. The interface temperature increased from 276.8℃ to 366.7℃, while the tensile-shear load was increased from 738.33 N to 2553.33 N when the clamping force was increased from 1100 N to 2200 N.

It is also evident from the observation that the interface temperature increases with the increasing value of bonding ratio as shown in Figure 9 (a). The mean values of interface temperatures for a particular set of levels also increased with the mean value of bonding ratios for that particular set of level data as shown in Figure 9 (b).

The bonding ratio represents the part of the area under sonotrode tip which converts into the actual weld area. If the value of the bonding ratio is on the higher side, it implies that more deformable area will be available for the heat flux due to deformation. A parallel relative motion is produced in the contacting surfaces due to the simultaneous action of the ultrasonic vibration. The higher value of clamping force brings the contacting surfaces in close contact with each other, thereby, providing more area for action. As a result, heat generation at the overlapped areas increases which reduces the resistive strength of the oxide and other contaminants against the progressive plastic deformation as well as shearing of the surface asperities. Ultimately, the nascent metallic surfaces start forming micro-welds. The experimental results also corroborate the simulated results.

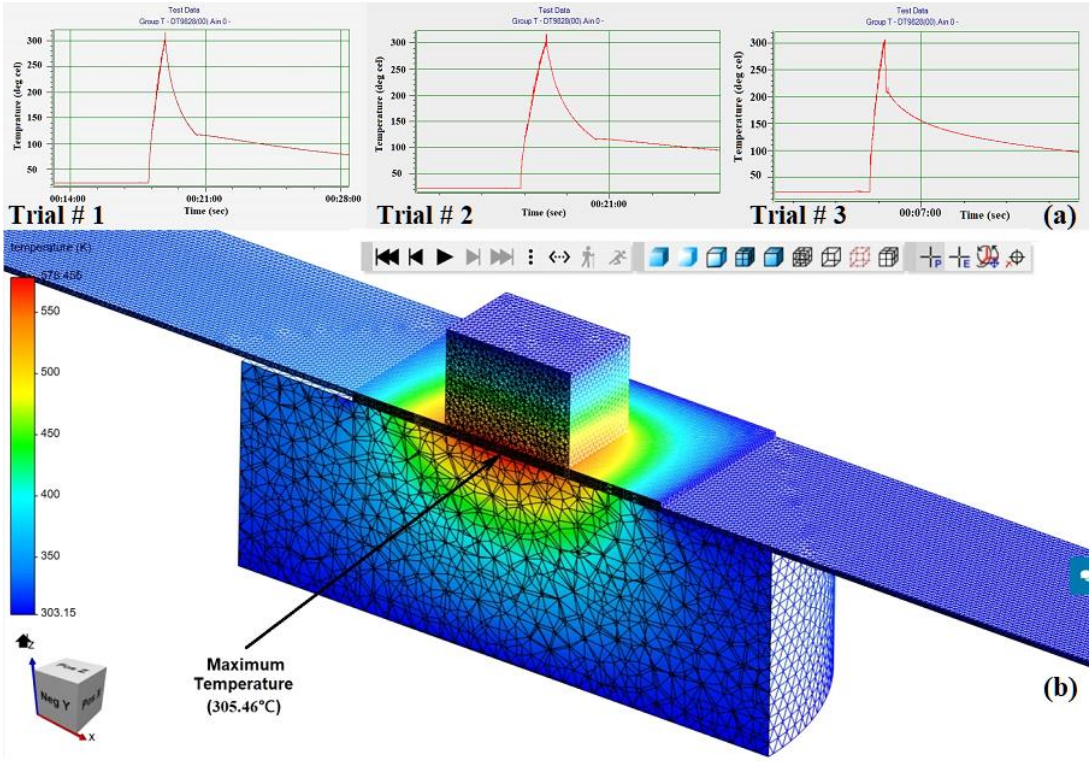


Figure 7. (a) Temperature profile of experimental results and (b) simulation result of run 14

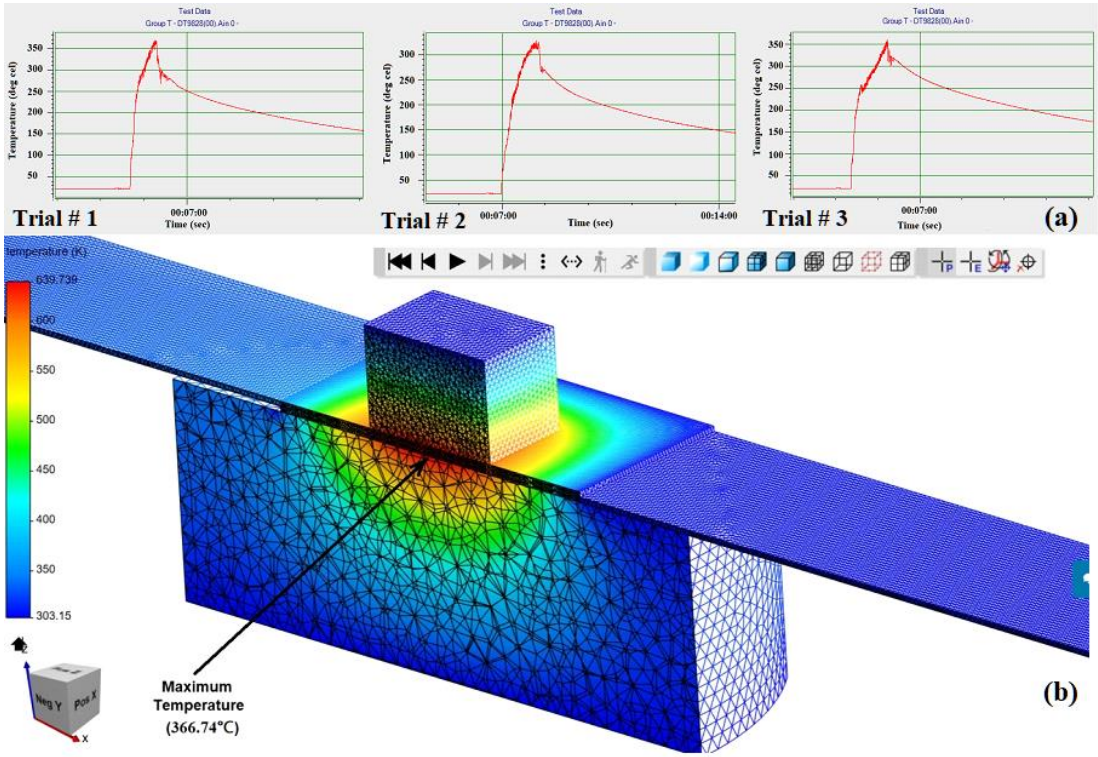


Figure 8. (a) Temperature profile of experimental results and (b) simulation result of run 10

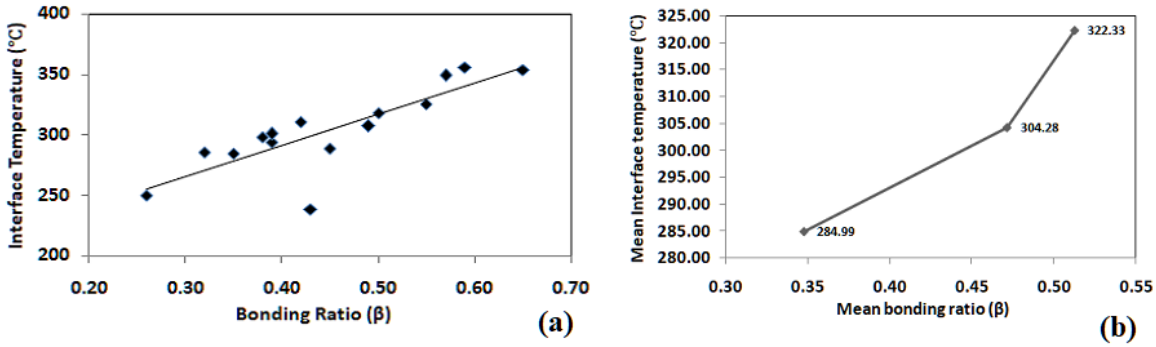


Figure 9. Effect of bonding ration on simulated results of interface temperature: (a) increasing trend in all the simulated runs and (b) increasing trend of mean interface temperature from low level to high level of parameters

Prediction of Thermal Profile in Weld Zone

The thermal analysis of the USW process can be utilized to predict the temperature at different zones in the weld interface as well as in different parts of the assembly such as sonotrode and anvil. Run number 10 having the highest temperature value and joint strength has been used for prediction analysis. Figure 10 shows the uniformly distributed temperature zones in the upper sheet, lower sheet, sonotrode tip, and anvil.

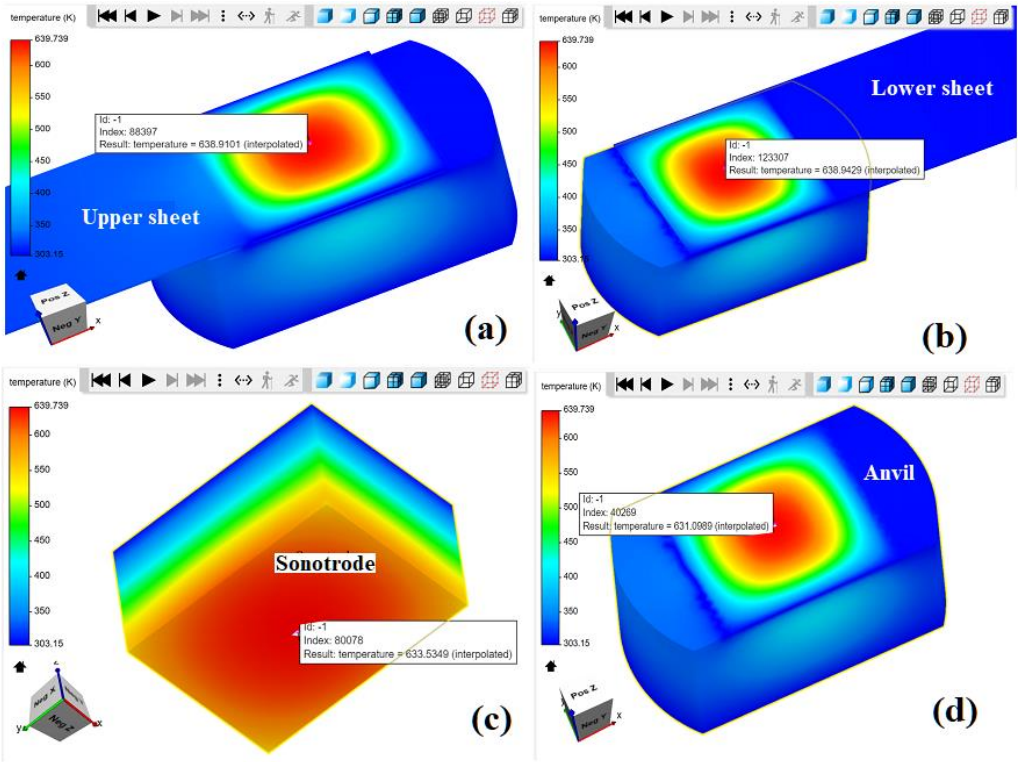


Figure 10. Temperature contour profile in: (a) upper sheet, (b) lower sheet and (c) sonotrode tip and (d) anvil

The temperature at the centre of the zone is also shown. The red colour zones indicated the highest temperature near the weld zone which gradually changed to blue colour away from it indicating the lowest temperature zone. The value of peak temperature in this zone is 366.7℃. This is a sufficiently high temperature to mobilize the considerable interface activities. Besides, significant plastic deformation having taken place, the oxides layer also gets burst and pure metallic adhesion takes place [30]. The highest temperature is experienced by the area under the sonotrode tip as shown in Figure 10 (a) and (b). Therefore, the top part and the interface of both the sheets go through severe plastic deformation, and the material starts yielding. The surrounding areas, however, are not susceptible to much yielding as is evident from the blue colour of the weld metal sheets. The temperatures towards the extremes of sonotrode and anvil are very less as compared to the weld zone due to their low thermal conductivity as well as low contact thermal conductance relative to the weld material at contacting surfaces. But the surfaces of the sonotrode, and the anvil which come into contact with the parent metal get instantaneous heat, and the temperature at the centre of these surfaces reaches at 360.7℃ and 357.3℃ respectively as shown in Figure 10 (c) and (d).

Estimation of TMAZ and HAZ

These temperature predictions corroborate the existence of TMAZ and HAZ in the vicinity of the weld zone. To explore further thermal activities, the temperatures at the end of the weld cycle at the weld zone were observed. Figure 11 shows the variation of the temperature at the weld interface starting from the centre line below the sonotrode tip and in its surroundings. There is a continuous variation in the colour of the interface both in X and Y directions ranging from the hottest (red colour) to the coolest (blue colour) temperature. The central part has maximum temperature where the effects of clamping force and ultrasonic vibrations are concentrated. The interface in this region undergoes severe plastic deformation and heat generation due to friction, thus, has the highest temperature. The oxide layer breaks and asperities reduce in this region to bring nascent metals to come in close contact to initiate micro-bonding.

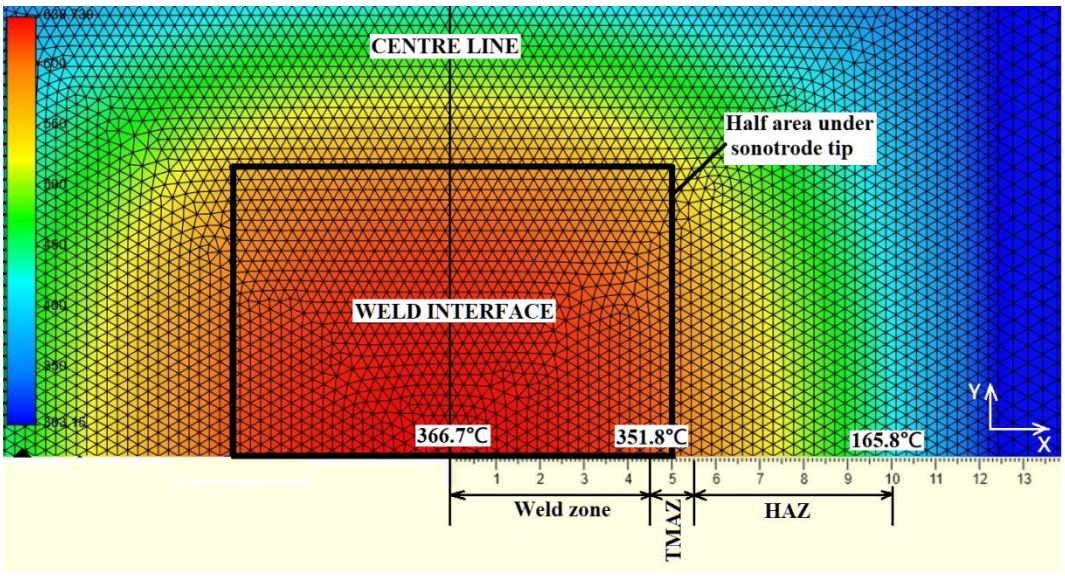


Figure 11. Temperature variation in weld interface along X axis

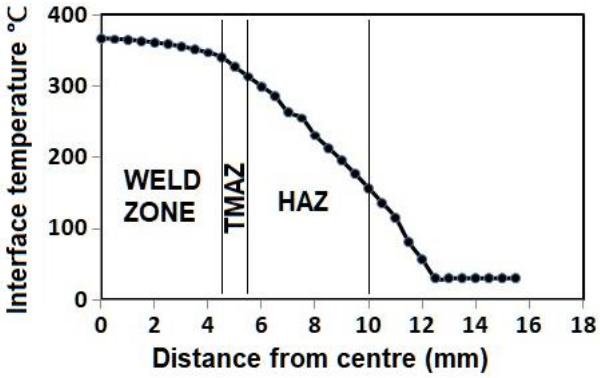


Figure 12. Different zones in weld interface

Plotting the interface temperature along the X-axis with respect to the distance from the centre (Figure 12) reveals more interesting information. It is observed that the maximum interface temperature at the end of the weld cycle is almost constant up to a certain distance along X-axis. The temperature at the centre line is 366.7℃ which decreases to 351.8℃ at 4.5 mm from centre line under sonotrode tip. It is the area that experiences severe plastic deformation. The micro-bonding initiates in this region. This area is known as the 'deformation zone' or 'weld zone'. There is a sharp fall in the temperature afterward. For about 1 mm, the temperature falls from 340.4℃ to 318.0℃. This zone also experiences metal softening due to plastic deformation and frictional heat. This is called the 'friction zone' and may be considered as TMAZ. It spreads beyond the boundary of the sonotrode tip. The recrystallization temperature is usually 0.3 to 0.6 times the melting temperature [24]. The temperature in TMAZ is within this range. The temperature continues to fall further and reaches 165.8℃ at a 10 mm distance from the centre. This is the 'non-contact zone' where no bonding takes place but some microstructural changes may be observed due to heat. This may be treated as HAZ. The boundary of the green colour in Figure 10 represents the end of HAZ.

Similarly, Figure 13 shows the thermal profile of sonotrode in an upward direction at the end of weld time. The colour distribution corresponding to the temperature varies from hottest (red colour, 352.42℃) at the bottom to the coolest (blue colour, 30.15℃) at the top within a height of 7 mm. It is obvious that the span of TMAZ is nil and HAZ is spread in a very small area in comparison to the weld metal. Very low thermal conductivity of sonotrode material and exposure of all the surfaces except one to the atmospheric air is responsible for this effect. The temperature variation of the sonotrode tip along the Y-axis is plotted in Figure 14. The temperature consistently decreased while moving upward from the bottom of the sonotrode with a temperature gradient of 46℃/mm. Reduced heat loss from the sonotrode helps in retaining more heat at the interface so that interface is able to attain a higher value of maximum temperature.

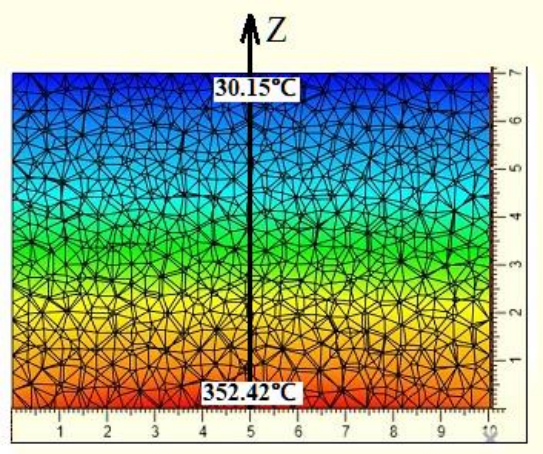


Figure 13. Temperature variation in sonotrode along Z axis

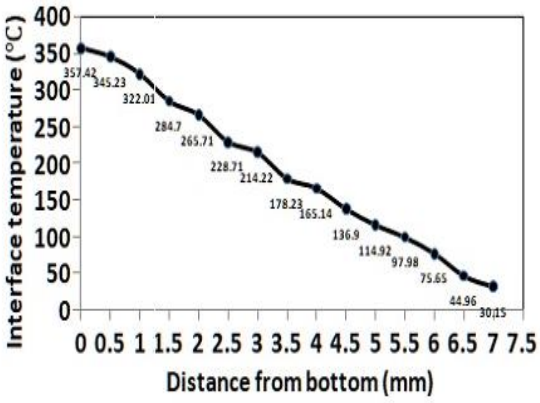


Figure 14. Plot between sonotrode temperature and distance along Z axis

Model Validation

A model is considered efficient if it gives comparable results. The temperature values measured by the thermocouple and those obtained by the simulation were compared and plotted in Figure 15. Both the results are in good agreement with each other with the maximum absolute error of 1.576%. It reflects that the developed FEA model can be efficiently utilized to predict the interface temperature and the boundary conditions presumed in the model may be treated as valid.

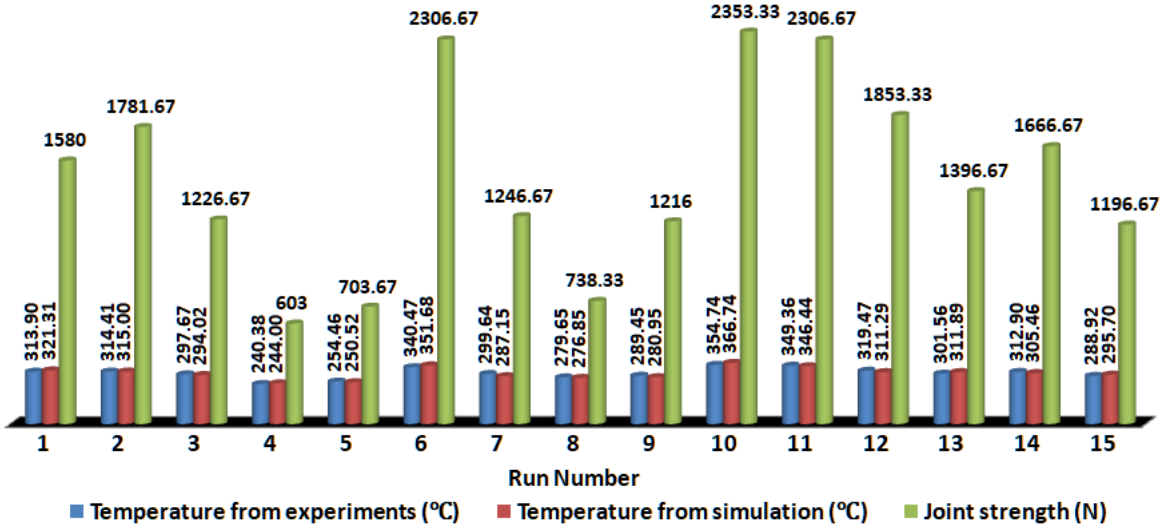


Figure 15. Comparison of interface temperature with joint strength

The ultimate tensile-shear load which is an indicative of the joint strength of the welded coupons was measured by pulling the joined weld coupons using UTM at a controlled specified speed. The optimal solution to achieve maximum joint strength was obtained at weld time = 1.2 sec, weld pressure = 0.36 MPa and vibration amplitude = 35 µm using Design-Expert software. The confirmatory tests performed at optimized parameter values resulted in the mean ultimate load of 2862.8 N which is about 74% of the strength of the base metal. Figure 12 also shows the joint strength along with the interface temperature. A correlation coefficient of 0.96 was calculated between interface temperature obtained by simulation and the joint strength to confirm their linear association. This value proves a strong correlation between the temperature developed at the interface and the weld strength. It signifies that the interface temperature may be considered to be an indicator of weld strength, and its value plays an important role in achieving stronger joints.

CONCLUSIONS

The finite element analysis of a 3-D model of USW is accomplished, and the model is utilized to predict the interface temperature. The following conclusions can be recapitulated:

i. The thermal model of the ultrasonically welded phosphor bronze sheets can effectively forecast the temperature results under various welding conditions. The peak temperature observed at the weld interface through

simulation is nearly 1/3rd of the melting point of phosphor bronze, hence, it is ascertained that USW is a solid-state joining technique.

ii. The maximum temperature observed at the centre of the weld interface is 366.74°C which diminishes continuously towards the extremes of the weld strips. Owing to the lower thermal conductivity of sonotrode material the maximum temperature obtained by it was lower in comparison to the maximum temperature of the weld interface.

iii. Clamping force and bonding ratio directly influence the interface temperature. Higher values of both the parameters produce a higher temperature.

iv. The three zones observed at the interface are weld zone, TMAZ, and HAZ. The TMAZ spreads up to 1 mm after the weld zone is finished and acted upon by severe acoustic softening. The HAZ is observed further up to 4.5 mm.

v. The model is validated with an error of 1.576% between the observed and predicted temperature results. Hence, the boundary conditions presumed in this model may be considered justified. A correlation coefficient of 0.96 established between the simulated temperature values, and the weld strength proves a strong linear relationship between them. Sufficiently strong joints were obtained at the optimum welding conditions with 74% joint efficiency.

Disclosure Statement

No potential conflict of interests was reported by the authors.

REFERENCES

[1] J. Yang and B. Cao, "Investigation of resistance heat assisted ultrasonic welding of 6061 aluminum alloys to pure copper," *Mater. Des.*, vol. 74, pp. 19–24, 2015, doi: 10.1016/j.matdes.2015.02.028.

[2] A. Das, "Joining Technologies for Automotive Battery Systems Manufacturing," *World Electr. Veh. J.*, vol. 9, no. 22, pp. 1–13, 2018, doi: 10.3390/wevj9020022.

[3] B. Mo, Z. Guo, Y. Li, Z. Huang, and G. Wang, "Mechanism of resistance microwelding of insulated copper wire to phosphor bronze sheet," *Mater. Trans.*, vol. 52, no. 6, pp. 1252–1258, 2011, doi: 10.2320/matertrans.M2011013.

[4] M. Weigl, M. Schmidt, E. Govekar, and A. Jeric, "Laser droplet generation : Application to droplet joining," *CIRP Ann. - Manuf. Technol.*, vol. 58, pp. 205–208, 2009, doi: 10.1016/j.cirp.2009.03.005.

[5] S. Consiglio, T. Fleschutz, G. Seliger, and J. Seutemann, "Development of a duothermal soldering process," *Ann. CIRP*, vol. 55, no. 1, pp. 1–4, 2006.

[6] K. Conway, P. P., Fu, E.K.Y., Williams, "Precision high temperature lead-free solder interconnections by means of high-energy droplet deposition techniques," *CIRP Ann.*, vol. 51, no. 1, pp. 177–180, 2002.

[7] C. Q. Zhang, J. D. Robson, O. Ciuca, and P. B. Prangnell, "Microstructural characterization and mechanical properties of high power ultrasonic spot welded aluminum alloy AA6111-TiAl6V4 dissimilar joints," *Mater. Charact.*, vol. 97, pp. 83–91, 2014, doi: 10.1016/j.matchar.2014.09.001.

[8] R. R. Dehoff and S. S. Babu, "Characterization of interfacial microstructures in 3003 aluminum alloy blocks fabricated by ultrasonic additive manufacturing," *Acta Mater.*, vol. 58, no. 13, pp. 4305–4315, 2010, doi: 10.1016/j.actamat.2010.03.006.

[9] B. Sanga, R. Wattal, and D. S. Nagesh, "Mechanism of joint formation and characteristics of interface in ultrasonic welding: Literature review," *Period. Eng. Nat. Sci.*, vol. 6, no. 1, 2018, doi: 10.21533/pen.v6i1.158.

[10] S. Elangovan, S. Venkateshwaran, and K. Prakasan, "Experimental investigations on optimization of ultrasonic welding parameters for copper to brass joints using response surface method and genetic algorithm," *Int. J. Adv. Eng. Res.*, pp. 55–64, 2012.

[11] D. Zhao, D. Ren, K. Zhao, S. Pan, and X. Guo, "Effect of welding parameters on tensile strength of ultrasonic spot welded joints of aluminum to steel – By experimentation and artificial neural network," *J. Manuf. Process.*, vol. 30, pp. 63–74, 2017, doi: 10.1016/j.jmapro.2017.08.009.

[12] Z. Ni, H. Zhao, P. Mi, and F. Ye, "Microstructure and mechanical performances of ultrasonic spot welded Al/Cu joints with Al 2219 alloy particle interlayer," *Mater. Des.*, vol. 92, pp. 779–786, 2016, doi: 10.1016/j.matdes.2015.12.132.

[13] M. P. Satpathy, S. B. Mishra, and S. K. Sahoo, "Ultrasonic spot welding of aluminum-copper dissimilar metals: A study on joint strength by experimentation and machine learning techniques," *J. Manuf. Process.*, vol. 33, no. April, pp. 96–110, 2018, doi: 10.1016/j.jmapro.2018.04.020.

[14] X. Zhong, J. Feng, and S. Yao, "Temperature field modeling and experimental study on ultrasonic consolidation for Al-Ti foil †," *J. Mech. Sci. Technol.*, vol. 33, no. 6, pp. 2833–2840, 2019, doi: 10.1007/s12206-019-0530-8.

[15] K. S. Suresh, M. R. Rani, K. Prakasan, and R. Rudramoorthy, "Modeling of temperature distribution in ultrasonic welding of thermoplastics for various joint designs," *J. Mater. Process. Technol.*, vol. 186, no. 1–3, pp. 138–146, 2007, doi: 10.1016/j.jmatprotec.2006.12.028.

[16] E. de Vries, "Mechanics and Mechanisms of Ultrasonic Metal Welding," *Ph.D. Thesis*, The Ohio State University, 2004.

[17] S. Elangovan, S. Semeer, and K. Prakasan, "Temperature and stress distribution in ultrasonic metal welding-An FEA-based study," *J. Mater. Process. Technol.*, vol. 209, no. 3, pp. 1143–1150, 2009, doi: 10.1016/j.jmatprotec.2008.03.032.

[18] E. Sooriyamoorthy, S. P. John Henry, and P. Kalakkath, "Experimental studies on optimization of process parameters and finite element analysis of temperature and stress distribution on joining of Al–Al and Al–Al2O3 using ultrasonic welding," *Int. J. Adv. Manuf. Technol.*, vol. 55, no. 5–8, pp. 631–640, 2010, doi: 10.1007/s00170-010-3059-7.

[19] K. K. Chen and Y. S. Zhang, "Numerical analysis of temperature distribution during ultrasonic welding process for dissimilar automotive alloys," *Sci. Technol. Weld. Join.*, vol. 20, no. 6, pp. 522–531, 2015, doi: 10.1179/1362171815Y.0000000022.

[20] P. Jedrasiak, H. R. Shercliff, Y. C. Chen, L. Wang, P. Prangnell, and J. Robson, "Modeling of the thermal field in dissimilar alloy ultrasonic welding," *J. Mater. Eng. Perform.*, vol. 24, no. 2, pp. 799–807, 2015, doi: 10.1007/s11665-014-1342-8.

[21] P. Jedrasiak and H. R. Shercliff, "Finite element analysis of heat generation in dissimilar alloy ultrasonic welding," *Mater. Des.*, vol. 158, pp. 184–197, 2018, doi: 10.1016/j.matdes.2018.07.041.

[22] U. I. Chang and J. Frisch, "On optimization of some parameters in ultrasonic metal welding," *Weld. J.*, vol. 53, no. 1, pp. 24s-36s, 1974.

[23] M. Shakil, N. H. Tariq, M. Ahmad, M. A. Choudhary, J. I. Akhter, and S. S. Babu, "Effect of ultrasonic welding parameters on microstructure and mechanical properties of dissimilar joints," *Mater. Des.*, vol. 55, pp. 263–273, 2014, doi: 10.1016/j.matdes.2013.09.074.

[24] A. K. Bhargava, M. K. Banerjee, and M. National, *Chapter 09196 - 2.14 Heat-Treating Copper and Nickel Alloys*, vol. 2. Elsevier Ltd., 2017.

[25] M. Winter, "Benchmark and Validation of open source CFD codes, with focus on compressible and rotating capabilities, for integration on the SimScale platform," 2013.

[26] K. Chen and Y. Zhang, "Thermal-mechanical analysis of ultrasonic spot welding considering acoustic softening effect," *Procedia Eng.*, vol. 81, October, pp. 2117–2122, 2014, doi: 10.1016/j.proeng.2014.10.295.

[27] J. R. Davis and D. & Associates, *ASM Specialty Handbook- Copper and copper alloys*, 2001st ed. ASM International, Materials Park, OH 44073-0002, 2001.

[28] L. S. Fletcher, "Recent Developments in Contact Conductance Heat Transfer," vol. 110, November 1988, 2013.

[29] P. Jedrasiak *et al.*, "Thermal Modeling of Al-Al and Al-Steel Friction Stir Spot Welding," *J. Mater. Eng. Perform.*, 32, 2016, doi: 10.1007/s11665-016-2225-y.

[30] D. Bakavos and P. B. Prangnell, "Mechanisms of joint and microstructure formation in high power ultrasonic spot welding 6111 aluminium automotive sheet," *Mater. Sci. Eng. A*, vol. 527, no. 23, pp. 6320–6334, 2010, doi: 10.1016/j.msea.2010.06.038.

Hindawi
Journal of Mathematics
Volume 2021, Article ID 5595536, 8 pages
https://doi.org/10.1155/2021/5595536

Research Article

An Introduction to Signed Petri Net

Sangita Kansal and Payal Dabas

Department of Applied Mathematics, Delhi Technological University, Bawana Road, Rohini, Delhi 110042, India

Correspondence should be addressed to Payal Dabas; payal.dtu@gmail.com

Received 3 February 2021; Revised 15 May 2021; Accepted 31 May 2021; Published 16 June 2021

Academic Editor: Ding-Xuan Zhou

Copyright © 2021 Sangita Kansal and Payal Dabas. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, the notion of a signed graph and Petri net (PN) is combined to introduce a new concept of *Signed Petri net* (SiPN). The balancedness in the signed graph, formulated by Harary and Cartwright, can be used to study and analyze social networks. This notion of balancedness is extended to Petri net via SiPN as such a concept has not been defined for already existing Petri nets. A balanced SiPN has been defined and its characterization is given. An illustrative example is provided which utilizes the newly introduced concept of assignment of sign to places of SiPN to determine whether an online transaction initiated by a bank customer is denied or approved by the bank. Lastly, applications of SiPN in the message transmission system and production unit are discussed. These concepts clearly demonstrate the advantages of the proposed approach of SiPN.

1. Introduction

A Petri net was first introduced in Carl Adam Petri's [1] dissertation submitted in 1962. However, it is difficult to model complex processes using the classical definition of Petri net given by him, and thus, many extensions and restrictions of Petri net have been proposed [2–9]. In the field of graph theory, Harary [10] introduced a signed graph and gave the characterization for a balanced signed graph. This concept of the signed graph has been used in [11] for predicting positive and negative links in online social networks and for determining the stability of the social model using the degree of balanceness [12]. Although various applications of PN have been given in the literature [4, 13, 14], the notion of balanceness of the modeled system does not exist in Petri net theory. The authors introduced a modified PN model with negative tokens for automated reasoning and to represent time constraint in batch activity, respectively [15, 16].

These developments in the field of graph theory and Petri net theory in the form of balancedness and negative tokens, respectively, motivated us to bridge the gap between Petri net and signed graph and introduce a new concept of *Signed Petri net* (*SiPN*). A SiPN is nothing but a bipartite signed digraph with dynamics associated with it in the form of tokens in places. There are various concepts in the signed graph that help in modeling of social interactions and checking the stability of the system so modeled. However, due to nondynamic nature of the signed graph, the system needs to be modeled using multiple signed graphs. Hence, we were interested in introducing an extension of Petri net called as Signed Petri net which can represent all possible interactions between a group of individuals via single SiPN as opposed to multiple signed graphs. The SiPN can be used to represent various configurations of a signed graph by a change in the marking of SiPN due to firing of transitions. This shows the advantage of the proposed research over a signed graph. We cannot differentiate among the tokens existing in a place of a PN. In order to overcome this limitation, two types of tokens (positive and negative) exist in a SiPN making it easier to represent the resources/processes common to a place via different types of tokens. The positive and negative arcs which exist in a SiPN make it possible to extend balancedness notion of a signed graph to a Petri net which can be used to model social networks and study them using a single SiPN rather than multiple signed graph.

Thus, a SiPN is an extension of PN which adopts the characteristics of both signed graph and PN and has advantages over the both. Furthermore, an illustrative example is given to demonstrate the advantages of the proposed research in Section 2.2 which uses the newly introduced concept of assigning sign to places of a SiPN.

The rest of the paper is structured as follows. In Section 2, SiPN has been introduced along with its related terminology. A method to assign a sign to vertices of a SiPN has been given along with an illustrative example to show the advantage of using a SiPN. In Section 3, balanced SiPN is defined and its characterization is given. In Section 4, we apply the proposed SiPN to model a message transmission system and a production unit.

1.1. Signed Graph. A *signed graph* is an ordered pair $S = (G, \sigma)$, where $G = (V, E)$ is called the underlying graph of S and σ: $E \longrightarrow \{+, -\}$ is a function which assigns either a positive or negative sign to each edge.

A signed graph depicts situations or structures in which both, a relation and its opposite, may occur, e.g., "like" and "dislike" in social relations, a simple reply of either "Yes" or "No" in communication. A signed graph has been used extensively to represent the interpersonal relationship among the group of individuals, study international relations between nations, and study the stability of traffic control problem at an intersection.

1.2. Petri Net. The definition of Petri net which we shall adopt in the paper is the one given by Jensen [2].

A *Petri net* (PN) is a 5-tuple $N = (P, T, I^-, I^+, \mu_0)$, where

(1) P is the finite, nonempty set of places

(2) T is the finite, nonempty set of transitions

(3) $P \cap T = \varnothing$

(4) I^-, I^+: $(P \times T) \longrightarrow \mathbb{N}$, where $\mathbb{N}$ is the set of nonnegative integers, are called negative and positive incidence functions, respectively

(5) $\forall p \in P$ and $\exists t \in T$ such that $I^-(p,t) \neq 0$ or $I^+(p,t) \neq 0$, and $\forall t \in T, \exists p \in P$ such that $I^-(p,t) \neq 0$ or $I^+(p,t) \neq 0$

(6) μ_0: $P \longrightarrow \mathbb{N}$ is the initial marking which gives the initial distribution of tokens in places

An arbitrary distribution of tokens in the places is called a *marking* given by

$$\mu\colon P \longrightarrow \mathbb{N}. \tag{1}$$

$I^-(p,t)$ represents the number of arcs from place p to transition t and $I^+(p,t)$ represents the number of arcs from transition t to place p. The number of arcs from a place p (or transition t) to a transition t (or place p) can also be considered as the *weight* of the arc (p,t) (or (t,p)).

A Petri net without its initial marking is called a *Petri net structure*.

A Petri net with all arcs of weight one is called an *Ordinary Petri net*.

A pair of place p and transition t is called a self-loop if p is both an input and an output place of t. A Petri net without any self-loop is called a *Pure Petri net*.

A PN $N' = (P', T', I_1^-, I_1^+, \mu_0')$ is called a *sub-Petri net* of a PN $N = (P, T, I^-, I^+, \mu_0)$ if

(1) $P' \subseteq P$

(2) $T' \subseteq T$

(3) $I_1^+(p,t) \neq 0$ iff $I^+(p,t) \neq 0$ for $(p,t) \in (P' \times T')$

(4) $I_1^-(p,t) \neq 0$ iff $I^-(p,t) \neq 0$ for $(p,t) \in (P' \times T')$

(5) $\mu_0'(p) = \mu_0(p)\ \forall\ p \in P'$

2. Signed Petri Net

Let $N = (P, T, I^-, I^+, \mu_0)$ be a PN. The arc set of the Petri net N is defined as

$$E = \{(p,t)\colon I^-(p,t) > 0\} \cup \{(t,p)\colon I^+(p,t) > 0\}. \tag{2}$$

A *signed Petri net* (SiPN) is defined as a 3-tuple $N^* = (N', \sigma, \mu_0)$, where

(1) $N' = (P, T, I^-, I^+)$ is a Petri net structure.

(2) σ: $E \longrightarrow \{+, -\}$, where E is the arc set of N'. An edge is called a positive or negative edge, respectively, according to the sign + or − assigned to it using the function σ.

(3) $\mu_0 = (\mu_0^+, \mu_0^-)$ is the initial marking of SiPN, where

(a) μ_0^+: $P \longrightarrow \mathbb{N}$ gives the initial distribution of positive tokens in the places, called positive marking of SiPN.

(b) μ_0^-: $P \longrightarrow \mathbb{N}$ gives the initial distribution of negative tokens in the places, called negative marking of SiPN.

Thus, a marking in SiPN can be represented as a vector $\mu = (\mu^+, \mu^-)$ with $\mu^+, \mu^- \in \mathbb{N}^n, n = |P|$, such that $\mu(p_i) = (\mu^+(p_i), \mu^-(p_i))\ \forall\ p_i \in P$.

In the graphical representation of a SiPN, positive and negative arcs are represented by solid and dotted lines, respectively. A positive token is represented by a filled circle and a negative token by an open circle, as shown in Figure 1.

A SiPN is said to be *negative* if all of its arcs are negative in sign.

The *underlying graph* of a SiPN N^*, denoted by N^ν, is the graph where directions of the arcs of N^* have been removed.

A *Complete SiPN* is a SiPN where every place-transition pair is connected by a bidirectional arc (either positive or negative).

A *path (cycle)* in a SiPN N^* is the path (cycle) in its underlying graph N^ν.

The *sign of a path* in a SiPN N^* is defined as the product of sign of arcs on the path.

The *sign of a cycle* in a SiPN N^* is defined as the product of sign of arcs on the cycle. If this product is positive, the cycle is called a *positive cycle*, and a *negative cycle* if the product is negative.

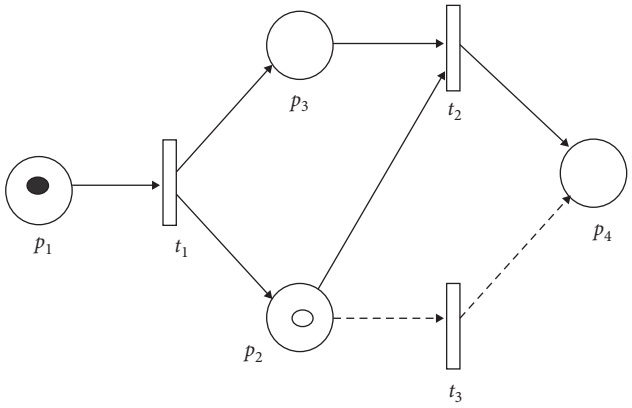


Figure 1: SiPN with initial marking ((1, 0, 0, 0), (0, 1, 0, 0)).

2.1. Execution Rules for Signed Petri Net. We know that the execution of a Petri net depends on the distribution of tokens in its places. The execution takes place by firing of a transition. A transition may fire if it is enabled.

A transition t in a SiPN N^* is *enabled* at a marking $\mu = (\mu^+, \mu^-)$ if

$$\begin{aligned} I^-(p,t) &\le \mu^+(p), \quad \forall p \in {}^\bullet t \text{ for which } \sigma(p,t) = +, \\ I^-(p,t) &\le \mu^-(p), \quad \forall p \in {}^\bullet t \text{ for which } \sigma(p,t) = -. \end{aligned} \tag{3}$$

An enabled transition t may *fire* at $\mu = (\mu^+, \mu^-)$ provided $\exists\, p_k \in t^\bullet$ such that

$$\sigma(t, p_k) = \begin{cases} +, & \text{if } \sigma(p,t) = + \; \forall p \in {}^\bullet t, \\ -, & \text{if } \sigma(p,t) = - \; \forall p \in {}^\bullet t, \\ + \text{ or } -, & \text{if } \sigma(p,t) = + \text{ for some } p \in {}^\bullet t \;\&\; - \text{ for some } p \in {}^\bullet t. \end{cases} \tag{4}$$

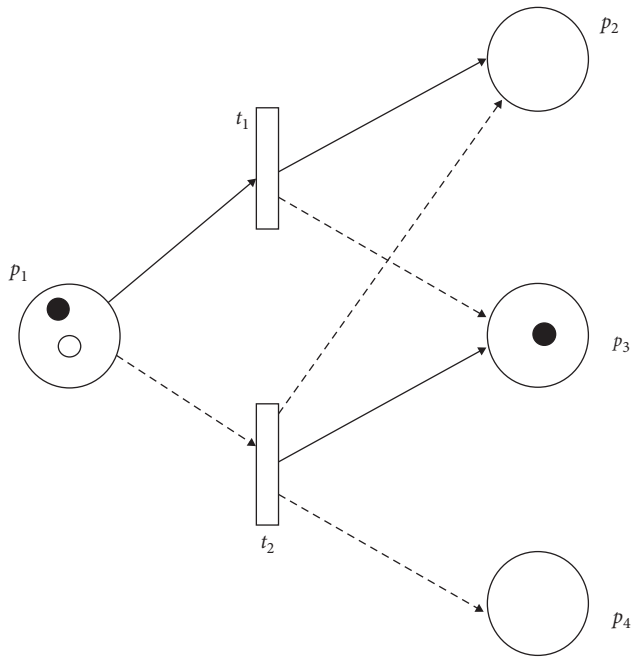


Figure 2: SiPN with $\mu_0 = ((1,0,1,0), (1,0,0,0))$.

After firing, it yields a new marking $\mu_1 = (\mu_1^+, \mu_1^-)$ given by the rule:

$$\begin{aligned} \mu_1^+(p) &= \mu^+(p) - I^-(p,t) + I^+(p,t), \quad \forall p \in P \text{ where } (p,t) \;\&\; (t,p) \text{are positive arcs, if exist,} \\ \mu_1^-(p) &= \mu^-(p) - I^-(p,t) + I^+(p,t), \quad \forall p \in P \text{ where } (p,t) \;\&\; (t,p) \text{ are negative arcs, if exist.} \end{aligned} \tag{5}$$

We say that μ_1 is reachable from μ and write $\mu \longrightarrow^t \mu_1$. We restrict the movement of positive (negative) tokens to positive (negative) arcs only.

Remark 1. A source transition is always enabled while a sink transition is never enabled in a SiPN.

In Figure 1, transitions t_1 and t_3 are enabled at marking $\mu = ((1,0,0,0), (0,1,0,0))$ and can fire. Let us look at the execution of a SiPN with the help of some more examples.

In Figure 2, t_1 and t_2 both are enabled at μ_0. Firing of t_1 at μ_0 yields a new marking $\mu_1 = ((0,1,1,0), (1,0,1,0))$, while firing of t_2 at μ_0 yields the marking $\mu_2 = ((1,0,2,0), (0,1,0,1))$. In Figure 3, t_1 is enabled, while t_2 is not. t_1 can fire at μ_0 to give a new marking $\mu_1 = ((0,0,1,0), (0,0,0,1))$.

2.2. Assignment of Sign to Vertices of a SiPN. The vertices in a SiPN can also be assigned sign. A transition is an assigned sign by the product of sign of arcs (incoming and outgoing) incident on it. In Figure 1, all transitions are positive in sign.

Places can be the assigned sign in one of the two ways:

(1) *With Respect to Arcs.* Sign is assigned to a place by taking the product of incident arcs (incoming and outgoing) on that place. In Figure 1, places p_1 and p_3 are positive in the sign, while p_2 and p_4 are negative in the sign.

(2) *With Respect to Marking.* Sign is assigned to a place by taking the product of sign of tokens in that place in the given marking. A place without token is considered to be positive. In Figure 1, places

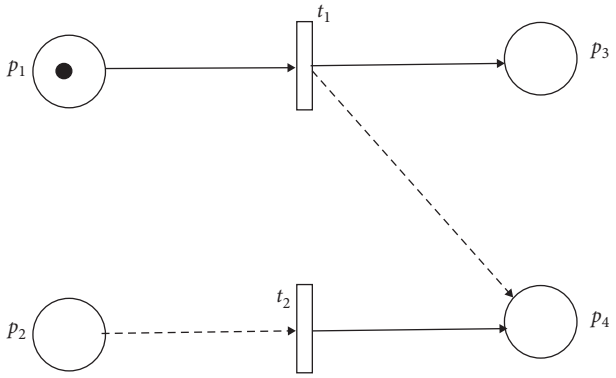


Figure 3: SiPN with $\mu_0 = ((1,0,0,0),(0,0,0,0))$.

p_1, p_3, and p_4 are positive in sign, while p_2 is negatively signed with respect to μ_0.

Remark 2. Assigning sign to places with respect to arcs does not utilize the most important characteristic of PN which is its dynamic behaviour. Hence, assigning a sign to places with respect to marking has been used throughout the paper.

An example is given which utilizes the concept of the place sign to determine whether an online transaction initiated by a bank customer is approved or denied by the bank. This transaction is based on the verification of one-time password (OTP) sent by the bank to the registered mobile number of the customer. This situation is modeled by a SiPN, as shown in Figure 4.

When a customer enters an OTP, firing of transition t_1 takes place. The number of times that t_1 fires is equal to the number of digits in the OTP. On firing of t_1, a positive token is generated in place p which represents the digit entered by the customer. This entered digit is then verified by comparing it with the corresponding digit of the OTP (sent by the bank), which exists in the bank system. If these two digits do not match, then the positive token is converted to a negative token, else it remains the same. On verification of the digit, firing of either t_2 or t_3 takes place depending on whether a positive or negative token is present in place p. After the firing of either t_2 or t_3, the sign of place q is checked. If the sign of place q with respect to this marking is negative, then the transaction is denied by the bank. On the contrary, if it is positive, the next digit of the OTP is checked in a similar manner until all the digits are exhausted.

Thus, we can conclude whether a transaction is denied or approved by the bank based on the sign of place q. If, at any marking, the sign of place q is negative, bank denies transaction to the customer. However, if, for all the markings with $\mu(q) \neq 0$, the sign of place q is positive, then the transaction is approved by the bank.

2.3. Reachability Tree of Signed Petri Net. The *reachability set* $R(N^*, \mu)$ of a SiPN N^* is the set of all markings of N^* reachable from μ.

A *reachability tree* represents the reachability set of a given SiPN. The reachability tree of the SiPN in Figure 1 is given in Figure 5.

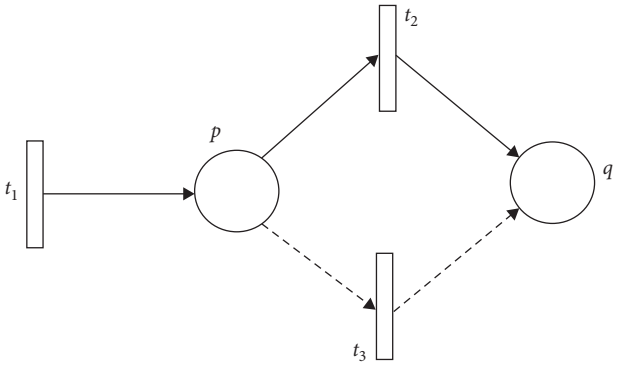


Figure 4: SiPN model for OTP verification.

3. Balanced Signed Petri Net

A SiPN N^* is said to be *balanced* if all of its cycles are positive in the underlying graph N^v. A SiPN without cycles in its underlying graph N^v is balanced.

Theorem 1. *A partition of the set of vertices $V = (P \cup T)$ in a SiPN N^* into two disjoint subsets V_1 and V_2, one of which may be empty such that all the arcs between vertices of the same subset are positive and all arcs between vertices of different subsets are negative, is possible in the following cases:*

(1) All the arcs in N^* are positive

(2) All the arcs in N^* are negative

(3) Any place in N^* has either positive or negative incoming as well as outgoing arcs

(4) Any transition in N^* has either positive or negative incoming as well as outgoing arcs

Proof

(1) In this case, we take $V_1 = V(= P \cup T)$ and $V_2 = \varnothing$.

(2) In this case, we begin with a place p_i in V and put it in subset V_1. Since only negative arcs exist among all the vertices, therefore, all the transitions connected to this place p_i must lie in the subset V_2. Now, the places connected to these transitions must lie in the subset V_1. Continuing in this manner, until all the vertices are put in either V_1 or in V_2, we obtain $V_1 = P$ and $V_2 = T$.

(3) By WLOG, we begin with a place with positive incoming and outgoing arcs and put it in the subset V_1. Since the transitions which are connected to this place are connected by positive arcs only, put these transitions in V_1 as well. Now, the transitions can have both types (positive and negative) of arcs incident on them. The places connected to these transitions via positive arcs lie in subset V_1 and those connected through negative arcs lie in subset V_2. Again these places are connected by only one kind of arcs so places in subset V_1 will have connections to transitions via positive arcs only and we put those transitions in subset V_1. The places placed in subset

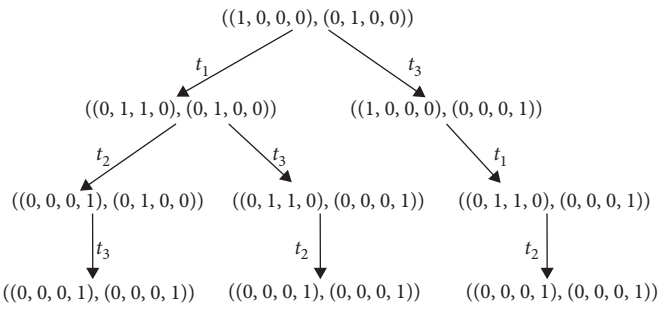


FIGURE 5: Reachability tree $R(N^*, \mu_0)$ of SiPN in Figure 1 with $\mu_0 = ((1, 0, 0, 0), (0, 1, 0, 0))$.

V_2 have negative arcs through which they are connected to other transitions, which then will also lie in subset V_1. We continue this process unless all vertices (places and transitions) are exhausted.

(4) In this case, we can proceed as in case 3. By WLOG, we begin with a transition having positive incoming and outgoing arcs. □

3.1. Observation. In addition to the four cases mentioned in Theorem 1, such a partition of vertices into two disjoint subsets in a pure SiPN N^* is possible in the following two cases also:

(1) Any place in a pure SiPN N^* has incoming and outgoing arcs of opposite signs (i.e., if incoming arcs are positive, then outgoing arcs are negative and vice versa)

(2) Any transition in a pure SiPN N^* has incoming and outgoing arcs of opposite signs

Corollary 1. *A SiPN N^* is negative if such a partition of the set of vertices, and V is a bipartite partition.*

Proof. If N^* is negative, then such a bipartite partition exists has been proved in Theorem 1 (2).

Conversely, suppose such a bipartite partition of V exists. Because of this partition, no arc exists within the set. Hence, no positive arcs exist resulting in an all negative SiPN.

Note: Theorems 2–4 and Corollary 2 have been formulated using the results given by Harary for a signed graph [10]. □

Theorem 2. *A complete SiPN N^* is balanced if its vertex set $V = (P \cup T)$ can be partitioned into two disjoint subsets V_1 and V_2, one of which may be empty such that all arcs between vertices of the same subset are positive and all arcs between vertices of different subsets are negative.*

Proof. Suppose such a partition of the vertex set $V = P \cup T$ exists in a complete SiPN. We need to show that all the cycles in N^ν are positive. Because of this partition, any cycle in N^ν always has even number of negative arcs if it exists; thus, the product of sign of arcs in the cycle is positive, making the SiPN balanced.

We prove the converse part by contradiction. Suppose a complete SiPN N^* is balanced and such a partition does not exist. Let $|V| > 3$ (for $|V| = 2, 3$ result is obvious). By WLOG, we assume that there exists at least one negative arc within a subset (say V_1) while trying to partition V. Let this negative arc exist between vertices v_i and v_j (place-transition pair). Since $|V| > 3$, there exist at least two more vertices v_k and v_l in V (place-transition pair). Since N^* is complete, v_k is connected to v_j and v_l is connected to v_i. Now, three cases arise:

(1) If $v_k, v_l \in V_1$, in this case, the arcs between v_i and v_l and between v_j and v_k are positive. The cycle of length four, so formed by vertices v_i, v_j, v_k, and v_l, is negative.

(2) If $v_k, v_l \in V_2$, in this case, the arcs between v_i and v_l and between v_j and v_k are negative. The cycle of length four, so formed by v_i, v_j, v_k, and v_l, is negative.

(3) If $v_k \in V_1$ and $v_l \in V_2$, here, the arc between v_i and v_l is negative while that between v_j and v_k is positive. Also, the arc between v_k and v_l is negative. The cycle of length four, so formed by v_i, v_j, v_k, and v_l, is negative.

If the partition does not exist on account of a positive arc across the subsets, then the result can be proved on similar lines.

In any of the cases, the negative cycle so formed will contradict the balanceness of N^*. Thus, our assumption is wrong, and there always exists such a partition. □

Corollary 2. *A sub-SiPN of a balanced SiPN is balanced.*

Proof. Since every cycle of a sub-SiPN is also a cycle of the given balanced SiPN and is therefore positive. Hence, sub-SiPN is also balanced. □

Theorem 3. *A SiPN N^* is balanced if, for each pair of distinct vertices v_1 and v_2, in the underlying graph N^ν, all paths joining v_1 and v_2 have the same sign.*

Proof. We are given a balanced SiPN. Consider any two paths A and B joining v_1 and v_2. If we remove any common arcs which exist in these paths, we get a collection of

arc-disjoint cycles. Each of these cycles consists of a subpath of A and a subpath of B. The cycle must be positive in sign which implies these subpaths must be of the same sign. Joining these subpaths with common arcs, we removed earlier leads to paths A and B having the same sign.

Conversely, we have all paths joining any two distinct vertices v_1 and v_2 in N^* of same sign. Hence, all cycles containing v_1 and v_2 must be positive. Since, v_1 and v_2 are arbitrary, all cycles in N^* are positive. Thus, N^* is balanced. □

Theorem 4. *A SiPN N^* is balanced if its set of vertices $V = (P \cup T)$ can be partitioned into two disjoint subsets V_1 and v_2; one of which may be empty such that all the arcs between vertices of the same subset are positive and all arcs between vertices of different subsets are negative.*

Proof. We are given a partition V_1 and v_2 of the set of vertices of N^*. We can extend N^* to a complete SiPN. Take a pair of nonconnected place and transition. If they lie in the same subset, join them with a positive bidirected arc, otherwise with a negative one. By Theorem 2, the complete SiPN so obtained is balanced, and by its Corollary 2, given SiPN is balanced.

Conversely, let N^* be balanced. Consider any nonadjacent pair (place-transition) of vertices v_1 and v_2. By Theorem 3, all the paths joining v_1 and v_2 have the same sign. Join v_1 and v_2 by a bidirected arc of the same sign as the sign of these paths. In this way, all the cycles thus introduced are positive, resulting in balanced SiPN. Once, we do this for all the nonadjacent vertices, we get a complete SiPN. Our result follows from Theorem 2. □

4. Applications of Signed Petri Net

4.1. Message Transmission System. When two parties need to communicate securely, i.e., they do not want a third party to listen in and they require a channel for secure communication. Secure communication is a way to transfer message from the source to the receiver so that it is not susceptible to eavesdropping or interception. We use the concept of SiPN which is introduced in this paper to model a message transmission system which is secure.

Consider a message which needs to be transferred from a source to a destination securely. In order to make the transmission secure, we divide our message into two parts, the first half of the message is represented by a positive token and the other half via a negative one. The places p_1 and p_6 are the source and the destination for the message, respectively. All other places are buffer for the message, which holds the message until it is ready to be transmitted, by firing of corresponding transitions. The transitions are events which transfer the message from one place to another. Note that $|{}^\bullet t| = |t^\bullet| = 1 \, \forall \, t \in T$.

In a SiPN, positive tokens move on positive arcs and negative tokens on negative arcs only. Therefore, we have a fixed path for the movement of both the tokens and no part of the message can move on the path of the other part of the message. Also, we restrict all transitions to have only arcs of one sign either positive or negative, i.e., for $p_i \in t^\bullet$,

$$\sigma((t, p_i)) = \begin{cases} +, & \text{if } \sigma((p,t)) = + \text{ for } p \in {}^\bullet t \\ -, & \text{if } \sigma((p,t)) = - \text{ for } p \in {}^\bullet t. \end{cases} \tag{6}$$

This restriction makes it impossible for any intermediate place/transition to have access to both parts of the message, thus avoiding leakage.

In Figure 6, a message is divided into two parts at place p_1 which is the source of the message. The positive token (which is the first half of the message) enables transition t_1 and the negative token (the other half of the message) enables transition t_2. When the transitions t_1 and t_2 fire, the message transmission begins. To reach the destination p_6, the positive token moves on path $t_1 p_4 t_4 p_5 t_6$ and the negative token on path $t_2 p_2 t_3 p_3 t_5$. These two parts of the message can then be combined to get the complete message at place p_6.

In this way, the message is transmitted from the source to the destination securely. The division of message into two parts and transfer of these parts via different paths reduce the chances of leakage of the message.

4.1.1. Extension of the Message Transmission System. In the extended model, places and transitions have usual meanings as in the model above, except for the transitions t_1 and t_8, which represent a process of dividing the incoming part of the message into two parts.

In Figure 7, a message is divided into two parts at place p_1 which is the source of the message. The positive token (which is the first half of the message) enables transition t_8 and the negative token (the other half of the message) enables transition t_1. When the transitions t_1 and t_8 fire, the parts of the message are further divided into two parts each. Thus, the original message is now divided into four parts. By WLOG, we assume that a positive token always represents the first half of the message or the first half of the part of the message. As the message transmission continues, the positive token (i.e., first half of the first part of the message) moves on the path $p_8 t_9 p_9 t_{10}$ and finally reaches the destination which is place p_{10}. Similarly, negative token (second half of the first part of the message) moves on the path $p_6 t_6 p_7 t_7$ and reaches the destination. In a similar fashion, other half of the original message is further divided into two parts by transition t_1 and reaches the destination p_{10}.

In this way, our message is moved by dividing it into four parts and transmitting these parts through different routes. This makes it difficult for anyone to access the whole message simultaneously, thereby making it secure. The model can be modified to divide the message into more parts so as to increase the level of secureness by dividing the message further, at the transitions which follow the transition t_1 and t_8.

4.2. Production Unit. We consider a production unit which produces and packages a product. The product is produced by first creating its parts, which are six in number separately

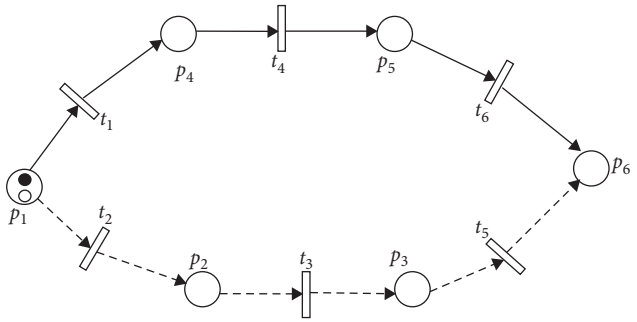


Figure 6: A message transmission system with message divided into two parts.

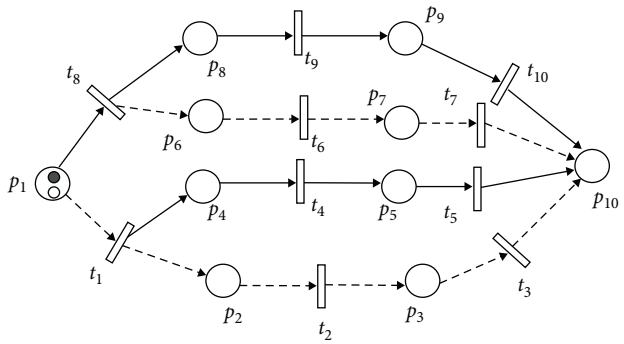


Figure 7: A message transmission system with the message divided into four parts.

and then assembling them together to get the final product in the assembling unit. In Figure 8, parts of the product are produced at place p_1. We assume that the parts which are heavier than a prescribed weight of W units are represented by a positive token and the remaining parts via a negative one. This is done so that the heavier parts move on the path $p_1t_2p_3t_4p_4$, which is suitable for transporting heavy parts, rather than on the path $p_1t_1p_2t_3p_4$ through which lighter parts are transported. All the transitions in the SiPN represent an event of the transfer of token (part) from one place to another. The places p_2 and p_3 are the test units which check the working of a part. Whenever a part of a product comes, it is tested for the required characteristics (depending on the product), and based on it, the part is segregated after being classified as "Tested Okay" or "not." Thus, the transitions t_4 and t_7 can fire accordingly. If the part is working correctly, it passes the test and moves to the assembling unit via firing of transition t_3 or t_4. On the contrary, if the part turns out to be defective, it is rejected and moved to place p_6 by firing of transitions t_6 or t_7. Whenever a part moves to the assembling unit or it is rejected, production of such a part should start again. Hence, we add arcs (t_3, p_1), (t_4, p_1); (t_6, p_1), (t_7, p_1).

If none of the six parts are rejected, the final product is assembled at place p_4. We represent this final product by a token whose sign is equal to the product of sign of all the tokens representing parts of the product (this sign is positive here). This final product is then moved to place p_5 for packaging by firing of transition t_5. It should be noted that the sign of arcs (p_4, t_5) and (t_5, p_5) is the same as the sign of token representing the final product.

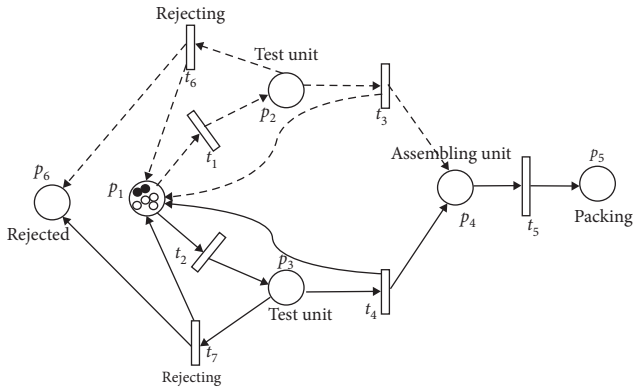


Figure 8: A production unit.

5. Conclusion and Scope

The concept of signed Petri net has been introduced along with the characterization of a balanced SiPN. An example which utilizes the newly introduced concept of assignment of sign to places of a SiPN is given which demonstrates the advantages of using a SiPN. The application of SiPN in the message transmission system and production unit is discussed.

The concepts of dynamical balancedness and domination have been developed by the authors, which we plan to introduce in our future work. These concepts do not exist for previously existing PN models and show how the proposed extension of SiPN is advantageous.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] C. Petri, *Kommunikation Mit Automaten,*Schriften des Institutes fur Instrumentelle Mathematik, Bonn, Germany, 1962.

[2] K. Jensen, "Coloured petri nets," *Petri Nets: Central Models and their Properties*, vol. 254, pp. 248–299, 1987.

[3] M. Marsan, *Stochastic Petri Nets: An Elementary Introduction: LNCS Advances in Petri Nets*, Springer, Berlin, Germany, 1989.

[4] T. Murata, "Petri nets: properties, analysis and applications," *Proceedings of the IEEE*, vol. 77, no. 4, pp. 541–580, 1989.

[5] W. Reisig, "Petri nets and algebraic specifications," *Theoretical Computer Science*, vol. 80, no. 1, pp. 1–34, 1991.

[6] R. Valk, *Object Petri Nets, LNCS 3098, Lecture on Concurrency and Petri Nets*, Springer, Berlin, Germany, 2004.

[7] J. Peterson, *Petri Net Theory and Modeling of Systems*, Prentice Hall-Inc, Hoboken, NJ, USA, 1981.
[8] S. Kansal, G. Singh, and M. Acharya, "On the Petri nets generating all the binary *n*-vectors," *Scientiae Mathematicae Japonicae*, vol. 71, no. 2, pp. 209–216, 2010.
[9] S. Kansal, G. Singh, and M. Acharya, "1-safe Petri nets generating every binary *n*-vector exactly once," *Scientiae Mathematicae Japonicae*, vol. 99, pp. 127–134, 2011.
[10] F. Harary, "On the notion of balance of a signed graph," *Michigan Mathematical Journal*, vol. 2, pp. 143–146, 1953.
[11] J. Leskovec, D. Huttenlocher, and J. Kleinberg, "Predicting positive and negative links in online social networks," in *Proceeding WWW'10 Proceedings of the 19th International Conference on World Wide Web*, pp. 641–650, Carolina, CA, USA, April 2010.
[12] D. Cartwright and F. Harary, "Structural balance: a generalization of heider's theory," *Psychological Review*, vol. 63, no. 5, pp. 277–293, 1956.
[13] W. M. P. van der Aalst, C. Stahl, and M. Westergaard, *Strategies for Modeling Complex Processes Using Colored Petri Nets, LNCS,Transactions on Petri Nets and Other Models of Concurrency VII*, Springer, Heidelberg, Germany, pp. 6–55, 2013.
[14] T. Brunsch, J. Raisch, L. Hardouin, O. Boutin, and O. Boutin, "Discrete-event systems in a dioid framework: modeling and analysis," in *Lecture Notes in Control and Information Sciences*, C. Seatzu, M. Silva, and J. van Schuppen, Eds., pp. 431–450, Springer, London, UK, 2013.
[15] T. Murata and H. Yamaguchi, "A Petri net with negative tokens and its application to automated reasoning," in *Proceedings of the 33rd Midwest Symposium on Circuits and Systems*, pp. 762–765, Calgary, Canada, August 1990.
[16] T. Brunsch, L. Hardouin, and J. Raisch, "Control of cyclically operated high-throughput screening systems," *IFAC Proceedings Volumes*, vol. 43, no. 12, pp. 167–172, 2010.

Applied Physics A (2021) 127:520
https://doi.org/10.1007/s00339-021-04652-0

Applied Physics A
Materials Science & Processing

Analytical modeling of dual-metal gate stack engineered junctionless accumulation-mode cylindrical surrounding gate (DMGSE-JAM-CSG) MOSFET

Sumedha Gupta[1] · Neeta Pandey[1] · R. S. Gupta[2]

Received: 3 April 2021 / Accepted: 1 June 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH, DE part of Springer Nature 2021

Abstract
This paper proposes a physics-based 2-D analytical model for a dual- material gate stack engineered junctionless accumulation-mode cylindrical surrounding gate (DMGSE-JAM-CSG) MOSFET. Analytical modeling is performed using 2-D Poisson's equation in cylindrical co-ordinate system based upon parabolic potential approximation. This analysis derives the expressions for center potential, electric field, subthreshold drain current, transconductance, output conductance and switching speed. It is seen that this device possesses enhanced drain current, higher transconductance and lower output conductance. The gate-induced drain leakage current for this device has also been investigated. The subthreshold slope of this device is approximately 71 mV/ decade and I_{on}/I_{off} ratio is also high, which leads to its usage for low power and high speed switching applications. The electrical characteristics and short channel effects of this device are also examined for different gate stack materials. It is observed that the device characteristics improve when permittivity of the gate stack is increased. Further, the results acquired using analytical modeling is mapped with the simulated data results to affirm and validate the device model structure. The simulation is implemented using ATLAS-3D device simulator.

Keywords Analytical model · Cylindrical surrounding gate · Junctionless accumulation mode · Parabolic potential approximation

1 Introduction

In conventional MOSFETs, when scaling of the device dimensions is performed, some short channel and hot carrier effects (SCEs and HCEs) arise [1, 2]. These effects result in the performance degradation of the device. Therefore, to suppress this problem, various topologies were presented such as double- gate (DG) MOSFETs, triple-gate MOSFET, cylindrical surrounding gate (CSG) MOSFET [3–5]. All these devices exhibit better performance in comparison with the conventional MOSFET. Due to superior gate control, CSG MOSFET offers high short channel immunity. But, they also possess a drawback of development of extremely—sharp source- drain junctions. So, to overcome this issue, junctionless transistor (JLT) was proposed [6]. This transistor has uniform doping all over; therefore, no junction is present in this and no doping concentration gradient is there. JLT offers improved electrical properties [7] and is also easy to fabricate. But because of the high doped channel, the mobility of the carriers gets degraded in the JLT. This problem of mobility degradation leads to lower drain current and lower transconductance. So, another structure, junctionless accumulation-mode (JAM) MOSFET [8, 9] was introduced. In JAM MOSFET, the channel region's doping is done slightly less than that of source and drain. On account of the higher doping present in the source and drain regions, it avoids high parasitic resistance; therefore, providing more conductivity and better characteristics than JLT [10]. JAM has reduced

✉ Sumedha Gupta
sumedhagpt5@gmail.com

Neeta Pandey
neetapandey@dce.ac.in

R. S. Gupta
rsgupta1943@gmail.com

[1] Department of Electronics and Communication Engineering, Delhi Technological University, Shahbad Daulatpur, Delhi 110042, India

[2] Department of Electronics and Communication Engineering, Maharaja Agrasen Institute of Technology, Rohini, Delhi 110086, India

Published online: 13 June 2021

SCEs but they are still not negligible, so, techniques of dual-metal and gate stack engineering are therefore added in this JAM MOSFET [11–15].

Architecture of dual-material gate employs two gates each with dissimilar work function such that the work function of the metal present at the source side is higher than the metal employed at the drain side. Dual gates offer excellent gate control which further leads to high drain current and high transconductance. Gate oxide engineering employs a high-k material layer stacked with the SiO_2 layer. This leads to reduction in leakage currents and is known as gate stack model. Now, using these techniques of dual-gates and a high-k gate stack, we have introduced a structure of dual-metal gate stack engineered JAM-CSG (DMGSE-JAM-CSG) MOSFET in 2020 [16]. Now, in this paper, 2-D analytical modeling of this device has been proposed to address center potential, electric field, subthreshold current, transconductance and various SCEs. Furthermore, the analysis is also done using different high-k gate stack materials. The exactness of this developed model is then established by comparing with the simulated outcomes.

2 Device description, calibration and simulation setup

Figure 1a represents the three-dimensional structure view and Fig. 1b represents two-dimensional cross-sectional view of DMGSE-JAM-CSG MOSFET. Let, the gate toward the source and drain side be represented as gate1 and gate2; with corresponding work functions as ϕ_1 and ϕ_2 (provided $\phi_1 > \phi_2$). In this structure, a high-k gate oxide layer is stacked with the SiO_2 layer. Analysis of the device will also be done by varying the gate stack material. Table 1 enlists various parameters utilized in carrying out the simulation. The device will also be investigated for various high-k gate stack materials. Table 2 lists different stack materials used, along with their dielectric constant values (k).

The simulation is performed using ATLAS-3D simulator with the help of the models mentioned in Table 3 [17]. Figure 2a and b illustrates the calibration results. For verification of the models, utilized for the device simulation, calibration has been performed with an experimental work of M. Najmzadeh et al. [18] and analytical work of J. Xiaoshi et al. [19]. Figure 2a indicates the drain current as a function of gate.

voltage and Fig. 2b indicates the drain current as a function of drain voltage. These graphs demonstrate the effectiveness of our model as the simulation data agrees well with the published experimental results and also with the analytical/ theoretical results.

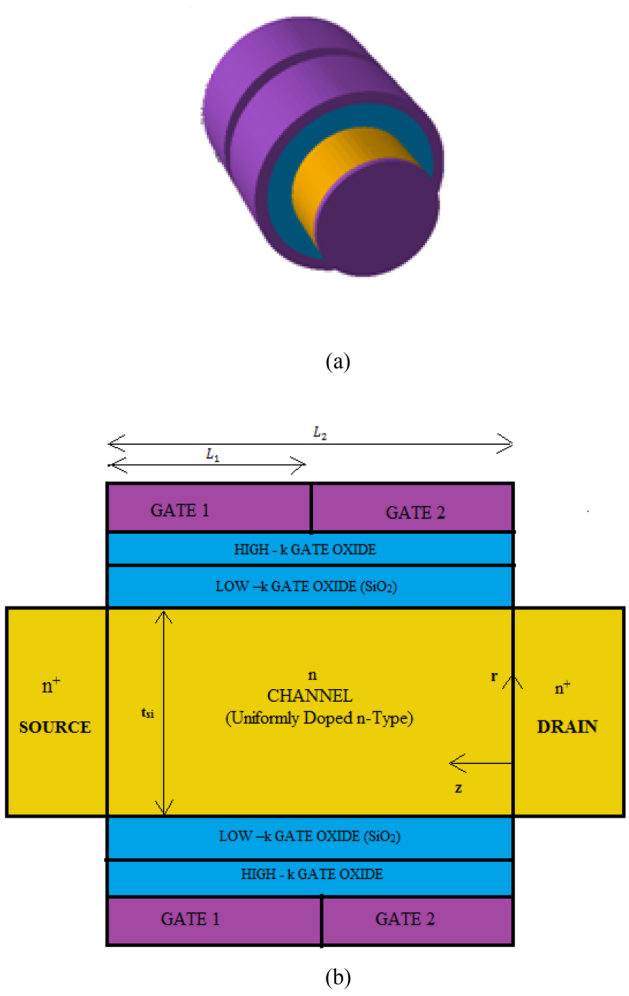


Fig. 1 **a** Three-dimensional structural view and. **b** 2-dimensional cross-sectional view of DMGSE-JAM-CSG MOSFET

Table 1 Parameters used

Parameter	DMGSE-JAM-CSG
Channel length	40 NM
Channel doping	$10^{18}/cm^3$
Source/drain doping	$10^{20}/cm^3$
Work function (ϕ_1)	5.0 eV
Work function (ϕ_2)	4.7 eV
Low-k gate oxide	SiO_2
Thickness of low-k gate oxide	1 nm
Thickness of high-k gate oxide	1 nm
Thickness of silicon pillar	20 nm
Permittivity of SiO_2	3.9

Table 2 High-k materials used

Material	Si_3N_4	Al_2O_3	Y_2O_3	Ta_2O_5	HfO_2	La_2O_3	TiO_2
k	7.5	10	15	22	25	30	40

Table 3 Physical models used for simulation

Physical Models	Description
Recombination Model	The Shockley–Read–Hall recombination model (with carrier lifetime 1×10^{-7} s) is included in simulation to incorporate minority recombination effects
Concentration-dependent mobility model	The concentration-dependent mobility model (CONMOB) is invoked to relate the low-field mobility at 300 K to the impurity concentration
Field-dependent mobility model	High-field mobility reduction model (FLDMOB) has been used in simulation
Band-to-band tunneling (BTBT) model	Standard band-to-band tunneling model has been used in the analysis to inculcate tunneling effect of the charge carriers
Statistics	Boltzmann Model because it considers the Carrier Statistics
Methods	The Newton and Gummel methods are simultaneously invoked for the numerical solution

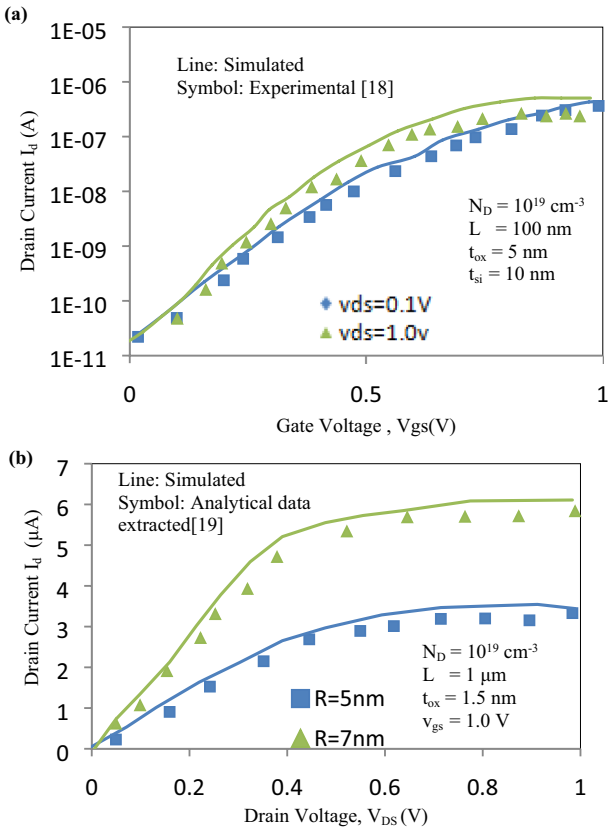


Fig. 2 **a** Calibration with M. Najmzadeh et al. [18] experimental work. **b** Calibration with J. Xiao-shi et al. [19] analytical data

3 Model derivation

Analytical analysis of DMGSE-JAM-CSG MOSFET is done using parabolic potential approximation while solving Poisson's equation [20–22] in cylindrical co-ordinates. The model has been developed within appropriate boundary conditions to obtain the center potential, electric field, subthreshold current, transconductance, output conductance and subthreshold slope.

3.1 Electrostatic potential

Considering that the channel region is fully depleted under subthreshold conditions, the impact of mobile charge carriers is negligible. The 2-D Poisson's equation in the cylindrical co-ordinates [20–22] is stated as follows:

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\psi_i(r,z)}{\partial r}\right)+\frac{\partial^2\psi_i(r,z)}{\partial z^2}=-\frac{qN_D}{\varepsilon_{si}},i=1,2 \tag{1}$$

$$\frac{\partial^2\psi_i(r,z)}{\partial r^2}+\frac{1}{r}\left(\frac{\partial\psi_i(r,z)}{\partial r}\right)+\frac{\partial^2\psi_i(r,z)}{\partial z^2}=-\frac{qN_D}{\varepsilon_{si}},$$

where $\psi(r,z)$ is potential distribution, r is radial co-ordinate, z is channel length co-ordinate, q is electronic charge, N_D is channel doping concentration and ε_{si} is the permittivity of SiO_2.

Solving Eq. (1), the potential along the radial or z-direction can be estimated using a parabolic profile function as follows:

$$\psi_j(r,z)=P_{j0}(z)+P_{j1}(z)r+P_{j2}(z)r^2,0\le z\le L \tag{2}$$

By substituting $r=0$ in Eq. (2), the center potential is obtained as follows:

$$\psi_j(r,z)|_{r=0}=\psi_j(0,z)=\psi_{cj}(z)=P_{j0}(z) \tag{3}$$

Equation (2) may be written for the two different materials as follows:

$$\psi_1(r,z) = P_{10}(z) + P_{11}(z)r + P_{12}(z)r^2, 0 \le L_1 \text{ (Region1)}$$
$$\psi_2(r,z) = P_{20}(z) + P_{21}(z)r + P_{22}(z)r^2, L_1 \le z \le L_2 \text{ (Region2)}, \quad (4)$$

where L_1 is the length of gate 1 and L_2 is the total channel length.

The Poisson's equation is interpreted distinguishably for both the regions (region1 and region 2) with the help of the following boundary conditions [20–22]:

(1). Channel's center potential is given by the following:

$$\psi(r,z)|_{r=0} = \psi_c(z) \quad (5)$$

(2). Surface potential of the channel toward source end is represented as

$$\psi(r, z=0) = \psi_s(r)|_{z=0} = \phi_{bi}, \quad (6)$$

where ϕ_{bi} is the built- in potential and is given as follows:

$$\phi_{bi} = V_T ln\left(\frac{N_D{}^+}{N_D}\right)$$

V_T is the thermal voltage and is given by, $V_T = \frac{k_B T}{q}$, where k_B is the Boltzmann's constant and T is the temperature. $N_D{}^+$ is the source/ drain doping and N_D is the channel doping concentration.

(3). Surface potential of channel toward drain end is given as follows:

$$\psi(r, z=L) = \psi_s(r)|_{z=L} = \phi_{\text{bi}} + V_{\text{ds}} \quad (7)$$

(4). Electric field in center of silicon substrate is zero, i.e.,

$$E = \frac{\partial\psi(r,z)}{\partial r}|_{r=0} = 0 \quad (8)$$

(5). Electric field at interface of silicon and stacked oxide layer is as follows:

$$\frac{\partial\psi(r,z)}{\partial r}|_{r=\frac{t_{si}}{2}} = \frac{\varepsilon_{\text{ox}}}{\varepsilon_{si}}\left[\frac{\psi_{\text{gs}} - \psi_{\text{si}}\left(r=\frac{t_{\text{si}}}{2}, z\right)}{\frac{t_{si}}{2} ln\left(1 + \frac{2t_{\text{oxeff}}}{t_{si}}\right)}\right], \quad (9)$$

whereψ_{gs} is the gate potential ($=V_{\text{gs}} - V_{\text{fb}}$), V_{gs} is the gate-source voltage, V_{fb} is flat band voltage ($=\phi_m - \phi_s$), ϕ_m and ϕ_s are work functions for metal and semiconductor, respectively,t_{si} is thickness of silicon region, t_{oxeff} is effective thickness of two oxide layers present in the gate stack and is given as follows:

$$t_{\text{oxeff}} = t_{\text{ox}} + \frac{\varepsilon_{\text{ox}.t_{\text{hk}}}}{\varepsilon_{\text{hk}}}$$

Here,t_{ox} and t_{hk} are the thicknesses of SiO_2 and HfO_2 layer, respectively, ε_{ox} and ε_{hk} are the permittivities of SiO_2 and HfO_2, respectively.

(6). At the interface of two distinct metals (metal 1 and metal 2), surface potential is continuous:

$$\psi_{s_1}(r, L_1)|_{r=\frac{t_{si}}{2}} = \psi_{s_2}(r, L_1)|_{r=\frac{t_{si}}{2}} \quad (10)$$

(7). Electric flux is also continuous at the interface of the metals:

$$\frac{\partial\phi_{s_1(R,z)}}{\partial z}|_{z=L_1} = \frac{\partial\phi_{s_2(R,z)}}{\partial z}|_{z=L_2} \quad (11)$$

On applying the above boundary conditions, the coefficients $P_{0i}(z)$,$P_{1i}(z)$ and $P_{2i}(z)(i = 0, 1, 2)$ we obtained are as follows:

$$P_{10}(z) = \psi_{c_1}(z)$$

$$P_{20}(z) = \psi_{c_2}(z)$$

$$P_{11}(z) = 0.$$
$$P_{21}(z) = 0$$

$$P_{12}(z) = \frac{C_{ox}}{\varepsilon_{si}.t_{si}}\left[\psi_{gs_1} - \psi_{si_1}\left(r=\frac{t_{si}}{2}, z\right)\right]$$
$$P_{22}(z) = \frac{C_{ox}}{\varepsilon_{si}.t_{si}}\left[\psi_{gs_2} - \psi_{si_2}\left(r=\frac{t_{si}}{2}, z\right)\right], \quad (12)$$

where C_{ox}, capacitance per unit area is given by the following:

$$C_{\text{ox}} = \frac{2\varepsilon_{\text{ox}}}{t_{si}\ln\left(1 + \frac{2t_{\text{oxeff}}}{t_{si}}\right)}$$

Now, using the above coefficient values, the center potential specified by Eq. (4) is rewritten as follows:

$$\psi_1(r,z) = \psi_{c_1}(z) + r^2\frac{C_{ox}}{\varepsilon_{si}.t_{si}}\left[\psi_{gs_1} - \psi_{si_1}\left(r=\frac{t_{si}}{2}, z\right)\right]$$
$$\psi_2(r,z) = \psi_{c_2}(z) + r^2\frac{C_{ox}}{\varepsilon_{si}.t_{si}}\left[\psi_{gs_2} - \psi_{si_2}\left(r=\frac{t_{si}}{2}, z\right)\right] \quad (13)$$

For $r = \frac{t_{si}}{2}$, $\psi\left(r=\frac{t_{si}}{2}, z\right) = \psi_{si}(r=\frac{t_{si}}{2}, z)$.

Solving (13) for ψ_{si_1} and ψ_{si_2} gives us

$$\psi_{si_1}\left(r=\frac{t_{si}}{2}, z\right) = \psi_{c_1(z)} + \left(\frac{t_{si}}{2}\right)^2 C_{ox}\left(\frac{\psi_{gs_1} - \psi_{si_1}\left(r=\frac{t_{si}}{2}, z\right)}{\varepsilon_{si}t_{si}}\right)$$
$$\psi_{si_2}\left(r=\frac{t_{si}}{2}, z\right) = \psi_{c_2(z)} + \left(\frac{t_{si}}{2}\right)^2 C_{ox}\left(\frac{\psi_{gs_2} - \psi_{si_2}\left(r=\frac{t_{si}}{2}, z\right)}{\varepsilon_{si}t_{si}}\right) \quad (14)$$

Rearranging the terms in (14) yields

$$\psi_{si_1}\left(r=\frac{t_{si}}{2},z\right)=\left[\frac{4\varepsilon_{si}\psi_{c_1(z)}+t_{si}C_{ox}\psi_{gs_1}}{4\varepsilon_{si}+t_{si}C_{ox}}\right]$$
$$\psi_{si_2}\left(r=\frac{t_{si}}{2},z\right)=\left[\frac{4\varepsilon_{si}\psi_{c_2(z)}+t_{si}C_{ox}\psi_{gs_2}}{4\varepsilon_{si}+t_{si}C_{ox}}\right] \quad (15)$$

The JLT works on bulk to source conduction mechanism, therefore, the center potential is to be calculated for $r=0$.

Center potential for two different metals is solved as:

$$\frac{\partial^2\psi_{c_1}(z)}{\partial z^2}=\frac{1}{\lambda^2}\left[\psi_{c_1(z)}-\eta_1\right],0\le z\le L_1,$$
$$\frac{\partial^2\psi_{c_2}(z)}{\partial z^2}=\frac{1}{\lambda^2}\left[\psi_{c_2(z)}-\eta_2\right],L_1\ \le z\le L_2\ , \quad (16)$$

where the characteristic length (λ), η_1 and η_2 are, respectively, given by

$$\frac{1}{\lambda^2}=\frac{16C_{ox}}{4\varepsilon_{si}t_{si}+{t_{si}}^2C_{ox}}$$

$$\eta_1=\psi_{gs_1}+\frac{qN_Dt_{si}}{4C_{ox}}+\frac{qN_D{t_{si}}^2}{16\varepsilon_{si}}$$

$$\eta_2=\psi_{gs_2}+\frac{qN_Dt_{si}}{4C_{ox}}+\frac{qN_D{t_{si}}^2}{16\varepsilon_{si}}$$

General solution of Eq. (15) comes out as:

$$\psi_{C1}(z)=Ae^{z/\lambda}+Be^{-z/\lambda}+\eta_1$$
$$\psi_{C2}(z)=Ce^{z/\lambda}+De^{-z/\lambda}+\eta_2, \quad (17)$$

where the coefficients A, B, C and D we obtained are given by,

$$A=\frac{e^{-L_1/\lambda}\left(\phi_{bi}-\eta_1\right)-V_{\mathrm{DS}}/2+\eta_1}{\left(e^{-L_1/\lambda}-e^{L_1/\lambda}\right)}$$
$$B=\frac{e^{L_1/\lambda}\left(\phi_{bi}-\eta_1\right)-V_{\mathrm{DS}}/2+\eta_1}{\left(e^{L_1/\lambda}-e^{-L_1/\lambda}\right)}$$
$$C=\frac{e^{-L_2/\lambda}\left(\frac{V_{\mathrm{DS}}}{2}-\eta_2\right)-e^{-L_1/\lambda}\left(\phi_{bi}+V_{\mathrm{DS}}-\eta_2\right)}{e^{(L_1-L_2)/\lambda}-e^{(L_2-L_1)/\lambda}} \quad (18)$$
$$D=\frac{e^{L_2/\lambda}\left(\frac{V_{\mathrm{DS}}}{2}-\eta_2\right)-e^{L_1/\lambda}\left(\phi_{bi}+V_{\mathrm{DS}}-\eta_2\right)}{e^{(L_2-L_1)/\lambda}-e^{(L_1-L_2)/\lambda}}$$

3.2 Electric field

Electric field distribution of this device along channel length can be found by taking differentiation of center potential (16).

$$E=-\frac{d\psi_{C1}(z)}{dz},0\le z\le L_1,$$
$$E=-\frac{d\psi_{C2}(z)}{dz},L_1\le\ z\le L_2 \quad (19)$$

3.3 Subthreshold drain current

It is the current between source and drain when the transistor is operating in subthreshold region [22]. It is given as:

$$I_d\left(V_{\mathrm{gs}},V_{\mathrm{ds}},z\right)=\frac{\left[2\pi N_D\mu\left\{1-e^{\frac{-(qV_{\mathrm{ds}}}{(k_BT}}\right\}\right]}{\int_0^{L_1+L_2}\frac{1}{\int_0^r r.e^{\left(\frac{q\psi(V_{gs},V_{\mathrm{ds}},z}{k_BT}\right)}dr}dz}, \quad (20)$$

where μ is carrier's mobility.

3.4 Subthreshold slope (SS)

It is change in gate voltage, V_{GS} for every one decade change in drain current, I_{DS}, i.e.,

$$\mathrm{SS}=\left[\frac{d\log(I_{DS})}{d(V_{\mathrm{GS}})}\right]^{-1}mV/\mathrm{decade} \quad (21)$$

3.5 *Transconductance (g_m) and output conductance (g_d)*

Transconductance (g_m) is defined as the rate of change in drain current with respect to gate voltage while keeping the drain voltage constant and is given by,

$$g_m=\left(\partial I_d/\partial V_{\mathrm{gs}}\right)|_{V_{\mathrm{ds=constant}}} \quad (22)$$

Output Conductance is rate of change in drain current with respect to drain voltage while keeping the gate voltage constant and is given by,

$$g_d=\left(\partial I_d/\partial V_{\mathrm{ds}}\right)|_{V_{\mathrm{gs=constant}}} \quad (23)$$

3.6 Switching speed (τ)

Switching speed of a MOSFET gives the measure of the rate of change of the output in response to changes at the input. It is given by,

$$\tau = \frac{C_G V_{DD}}{I_{ON}}, \tag{24}$$

where C_G is the gate capacitance, V_{DD} is the power supply voltage and I_{ON} is the ON current of the MOSFET.

4 Results and discussion

The simulation has been carried out on ATLAS-3D simulator and the acquired results are also verified with the proposed developed model. The device is analyzed for different high- k materials, i.e., $k = 7.5$, 15 and 22. This is done to discover the impact of k variation on behavior of the device. Figure 3 indicates the center potential of the device across the channel length for various stack materials ($k = 7.5$, 15 and 22) for 40 nm channel length. It is observed from the graph that as permittivity of the material is increased, center potential gets lowered down. This decrease in potential distribution reflects decrease in SCEs [23]. It is also noticed that the simulation results agree well with the analytical model. Due to the higher work function of the metal present near the source end, the charge carriers existing in the channel region experiences an acceleration force and the other gate material that is near the drain end possessing lower workfunction brings about screening, which results in the lowering down of DIBL and channel length modulation (CLM) effect. Thus, yields in more immunity toward SCEs and HCEs. Figure 4 displays graph indicating electric field across channel of length equal to 40 nm for numerous

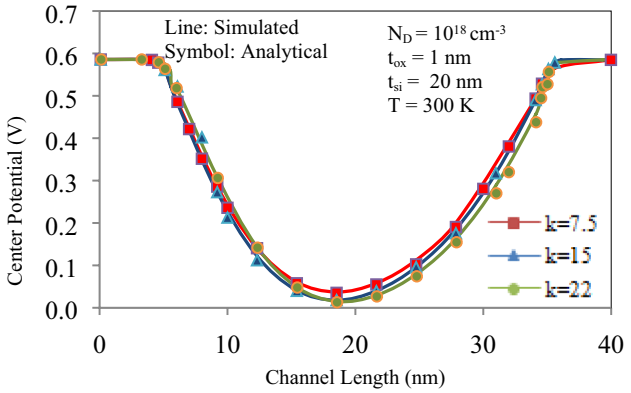


Fig. 3 Center potential over the channel length for various gate stack materials ($k = 7.5$, 15, 22)

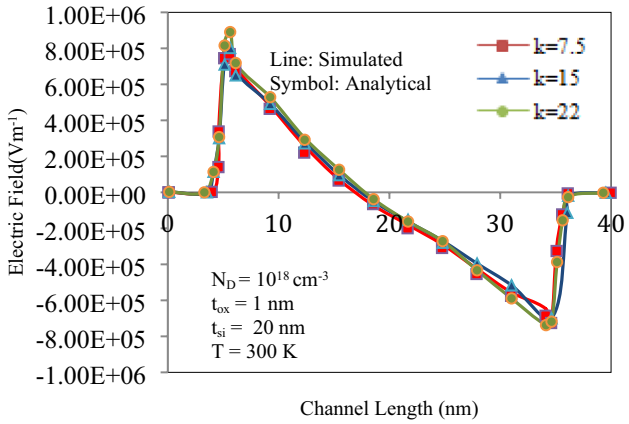


Fig. 4 Electric field over the channel length for various gate stack materials ($k = 7.5$, 15, 22)

materials ($k = 7.5$, 15, 22). It can be noticed from the figure that near the drain side, the peak electric field is lowered. This is due to the presence of lower work function at the drain side which results in reduced HCEs, lowered impact ionization and improved breakdown voltage. Also, the presence of the gate stack increases the gate controllability over the channel. From the figure, it is noted that as the material's permittivity increases, electric field toward the source end increases. Thus, the carrier transport efficiency is increased or there is rise in mean electron velocity. Further, there is close agreement between numerical and simulated results.

Leakage current has become a major issue in recent CMOS technology. Various machanisms exists to study this leakage current. Subthreshold leakage [24] is one among these and is a measure of the drain-to-source current that flows when MOSFET operates in the subthreshold region. Figure 5 shows graph of subthreshold current with gate voltage for various gate stack materials ($k = 7.5$, 15, 22). It is

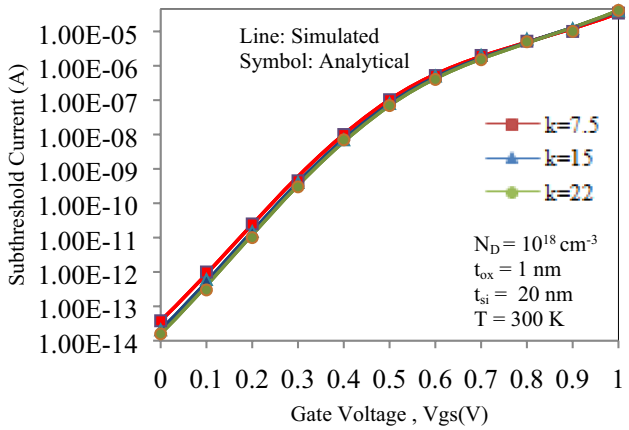


Fig. 5 Subthreshold current vs. gate voltage for various gate stack materials ($k = 7.5$, 15, 22)

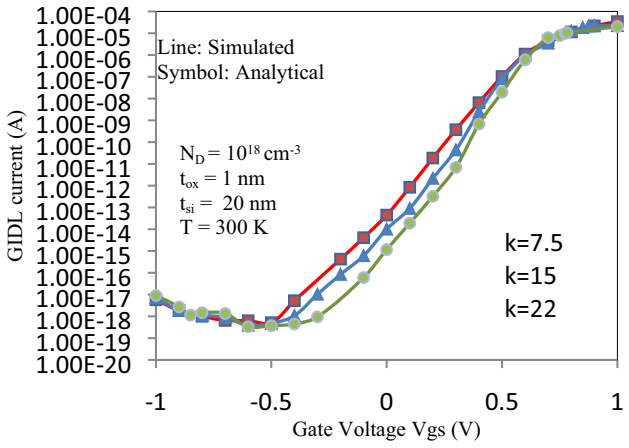


Fig. 6 GIDL current vs. gate voltage for various gate stack materials (k = 7.5, 15, 22)

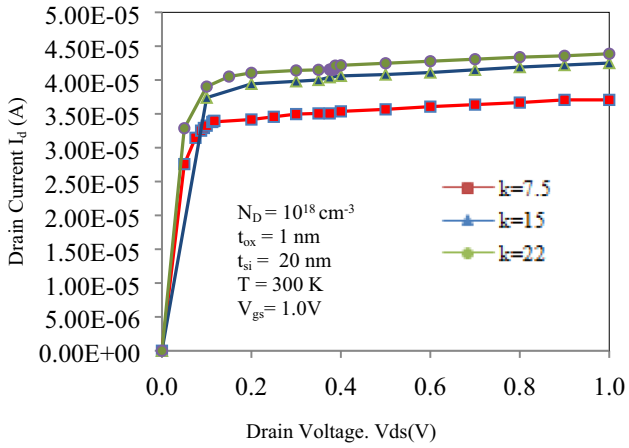


Fig. 8 Drain current vs. drain voltage at V_{gs} = 1 V for various gate stack materials (k = 7.5, 15, 22)

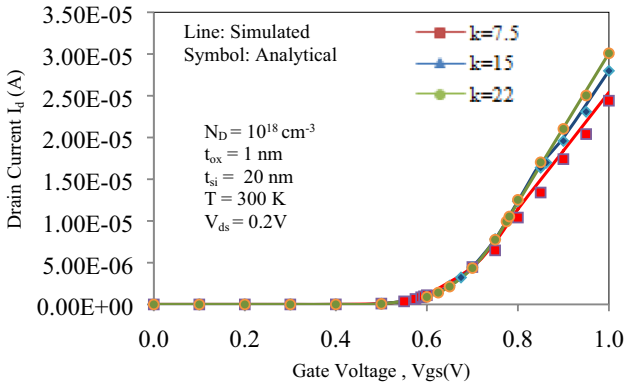


Fig. 7 Drain current vs. gate voltage at V_{ds} = 0.2 V for various gate stack materials (k = 7.5, 15, 22)

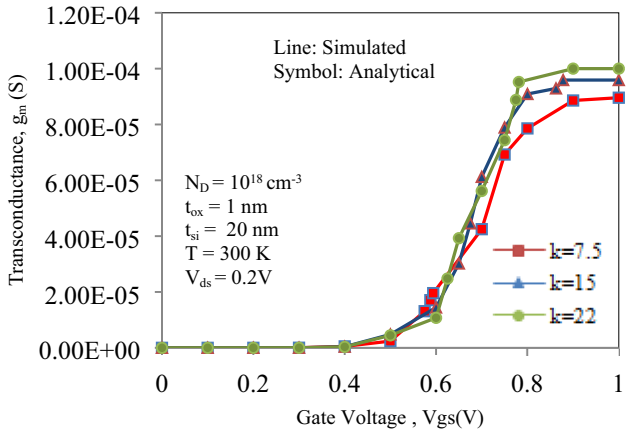


Fig. 9 Transconductance (g_m) with respect to gate voltage at Vds = 0.2 V for numerous gate stack materials (k = 7.5, 15, 22)

noticed that when dielectric constant is increased, subthreshold current reduces. This indicates improved device performance. Also, simulated data agrees well with the analytical data. Another mechanism for leakage current is gate-induced drain leakage (GIDL) [25] which is an attribute of presence of high electric field at the drain. This causes band-to-band tunneling current in the gate-to-drain overlap region. Figure 6 shows GIDL current (I_{gidl}) w.r.t. gate voltage for various dielectrics, k = 7.5, 15 and 22. It is seen that the value of I_{gidl} s quite low and its value reduces with increasing dielectric constant. At V_{gs} = 0.4 V, I_{gidl} reduces from 5.23E-18A to 4.45E-19A when k moves from 7.5 to 22. The reduced subthreshold current and GIDL current implies lowering of SCEs and marking our device to be more efficient.

Figure 7 presents graph displaying drain current against gate voltage at drain voltage, V_{ds} = 0.2 for different materials [26, 27]. It is noticed from this figure that I_{on} is much higher due to existence of gate stack architecture It is seen that the drain current rises with increment in dielectric constant. This effect is due to higher inversion charge resulting from increasing dielectric constant It is also seen that the simulation results correspond well with the analytical results, thus verifying our model. Figure 8 depicts the graph displaying drain current as a function of the drain voltage at V_{gs} = 0.1 V for various stack materials (k = 7.5, 15, 22) for L = 40 nm. Drain currents for all the three gate stack materials are almost equal in the linear region, whereas, in the saturation region, device with higher k exhibits higher drain current. So, there is a rise in the drain current with increase in material's dielectric constant, the reason is same as discussed above. The simulated results match well with the analytical ones.

Transconductance (g_m) is a vital specification. It determines the potential of the device to drive the load. Another important parameter is output conductance (g_d). Both these terms are already defined above in Sect. 3.5. Voltage gain of

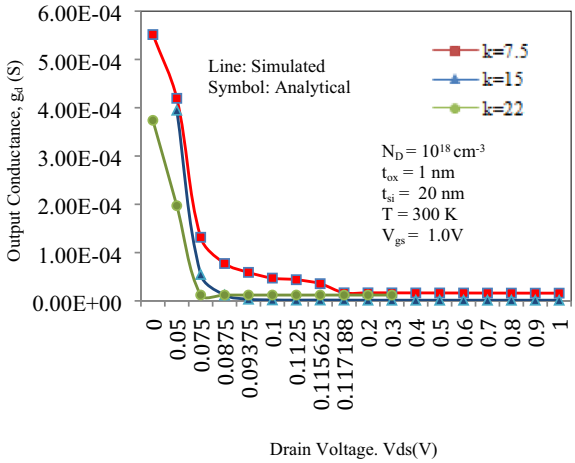


Fig. 10 Output Conductance (g_d) with drain voltage, Vds at Vgs = 1.0 V for different gate stack materials (k = 7.5, 15, 22)

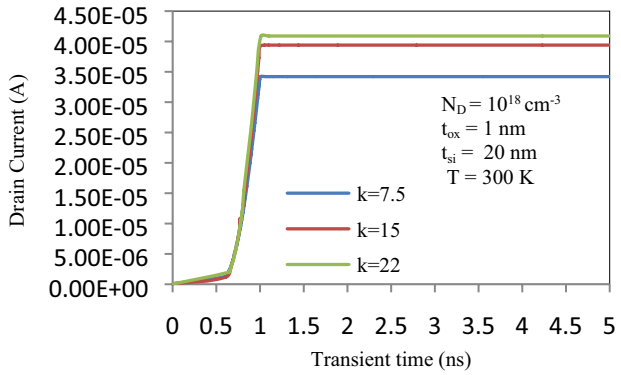


Fig. 11 Transient time response of the device for k = 7.5, 15 and 22

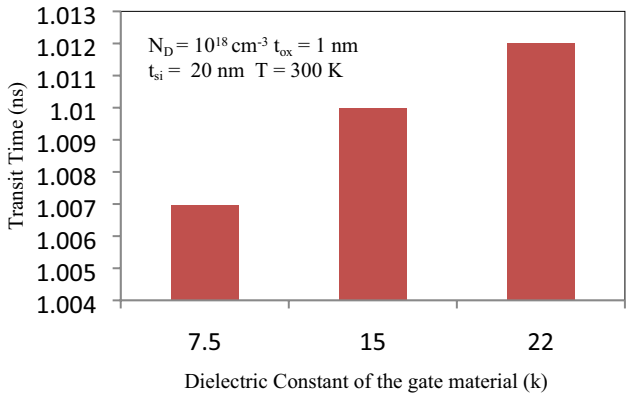


Fig. 12 Switching time of DMGSE-JAM-CSG MOSFET for the various gate stack materials used

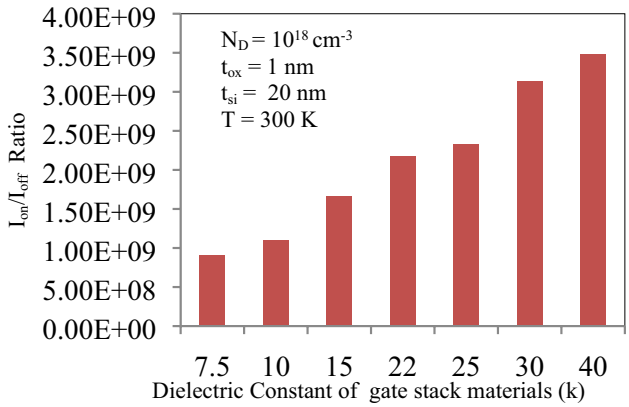


Fig. 13 I_{on}/I_{off} ratio with the dielectric constant of the different materials

MOSFET is defined as the ratio of the above two parameters, i.e., $A_v = (g_m/g_d)$. This means for higher gain, transconductance should be higher and output conductance should be lower. Figure 9 depicts the transconductance (g_m) against the gate voltage (V_{gs}) for several gate stack materials. As seen from the figure that transconductance is higher for k = 22 as compared to k = 7.5 and 15. So, for the higher dielectric constant material, transconductance increases. This indicates its suitability for high frequency and high gain amplifier applications.

Further, Fig. 10 indicates output conductance (g_d) with drain voltage (V_{ds}) for several stack materials (k = 7.5, 15, 22). It can be noted from Fig. 10 that as dielectric constant increases, output conductance reduces. Thus, resulting in increased voltage gain of the device. Also, increase in the dielectric constant of the stack will lead to increase in the gain of the device. Also, early voltage which is given as, $V_{EA} = (I_d/g_d)$ will also become high due to this lowered output conductance.

Another important criterion for verifying the device's performance is its switching speed as it provides an estimation of the delay experienced by the output in responding the input change. Using Eq. (23), switching speed is also computed for this device, i.e., DMGSE-JAM-CSG MOSFET for dielectrics, k = 7.5, 15 and 22. Figure 11 shows the transient time response [28] of this device i.e., DMGSE-JAM-CSG MOSFET for various dielectrics, k = 7.5, 15 and 22. From the graph, it is seen that when the device is turned ON, initially the drain current rises and after some time, it attains a saturated value (ON current). The time at which it attains this ON current is the switching time of the device. Figure 12 shows the variation of switching time with dielectric constant. It is observed that the device is switched ON approximately at the same time irrespective of the dielectric constant. The switching time is nearly at 1 ns which authenticates that our device is useful for faster switching applications.

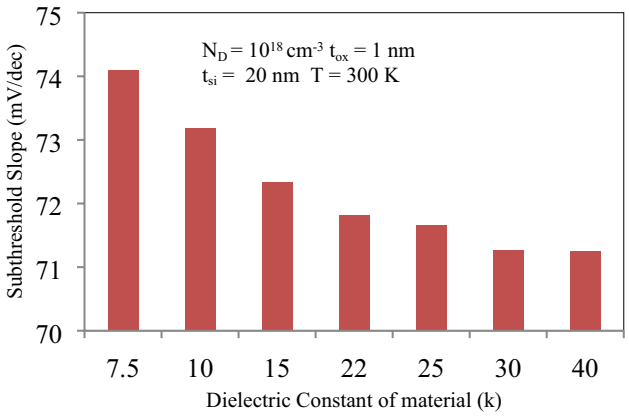


Fig. 14 Subthreshold slope for the various gate stack materials used

Table 4 Comparison with analogous devices

Performance Parameter		I_{on}/I_{off}	Peak g_m (S)
JL-CSG	$K=3.9$	3.85E+06	2.13E-05
	$K=10$	1.33E+07	2.03E-05
	$K=25$	7.89E+07	1.98E-05
JAM-CSG	$K=3.9$	7.43E+07	4.27E-05
	$K=10$	1.04E+08	3.96E-05
	$K=25$	2.00E+08	3.80E-05
DMGSE-JAM-CSG	$K=3.9$	9.00E+08	8.96E-05
	$K=10$	1.09E+09	8.88E-05
	$K=25$	2.33E+09	9.73E-05

Also, the operating speed can also be analyzed by computing the I_{on}/I_{off} ratio for this device. Figure 13 displays I_{on}/I_{off} ratio of this device when distinct gate stack materials are used. This figure shows that the device possesses very high and improved I_{on}/I_{off} ratio with increase in material's dielectric constant. Numerically, the I_{on}/I_{off} ratio improves from 9.00E+08 ($k=7.5$) to 3.48E+09 ($k=40$) indicating its better switching capability. This improvement is caused by the elevated vertical effect of the fringing electric field. Subthreshold slope (SS) is another parameter. Its value should be low, approximately equal to 60 mV/decade. Figure 14 depicts variation of subthreshold slope of the device with k. The value of SS is observed to be 74 mV/decade for $k=7.5$ which reaches approximately 71 mV/ decade for $k=30$. Thus, the device has better subthreshold slope with increasing dielectric constant. This indicates increased gate controllability over the channel. Hence, the overall device's performance is augmented with the increasing value of k of the gate stack material.

The performance parameters of DMGSE-JAM-CSG MOSFET are compared with two available analogous devices namely Junctionless (JL)-CSG MOSFET and JAM-CSG MOSFET [29]. The I_{on}/I_{off} ratio and peak transconductance values of these devices for various dielectric constants ($k=3.9$, 10 and 25) are listed in Table 4. The table shows an improvement in parameters for over JL-CSG MOSFET and JAM-CSG MOSFET for various dielectric constants. As can be seen from the table, there is a significant improvement in I_{on}/I_{off} ratio for DMGSE-JAM-CSG MOSFET as compared to other two devices (JL-CSG and JAM-CSG). It has increased by 12, 10 and 11 times in context to JAM-CSG MOSFET for the dielectrics, $k=3.9$, 10 and 25, respectively. Also, there is an increase in the value of transconductance (g_m) for the case of DMGSE-JAM-CSG MOSFET as compared to JL-CSG and JAM-CSG MOSFET. These values indicate the device is more robust and better suited for switching and high- frequency applications.

5 Conclusion

Utilizing 2-D Poisson's equation in cylindrical co-ordinates, an analytical model of DMGSE-JAM-CSG MOSFET is being introduced. Influence of dual-gate and stack engineering on JAM MOSFET for analog/ RF applications has been examined here for 40 nm channel length. The mathematical formulation for center potential, electric field, subthreshold current, transconductance and output conductance is put forward. The capability of the device has been observed by varying the gate stack materials. As, the dielectric constant increases, improvement in the device performance has also been observed. Results obtained reveals high I_{on}/I_{off} ratio, low subthreshold slope, high transconductance and low output conductance for the device when permittivity of the stack layer is increased. Also, the numerical results follow simulation results closely. The device DMGSE-JAM-CSG MOSFET offers improvement in SCEs and HCEs. Also, it is capable for rapid switching and high- frequency applications.

Acknowledgements Authors convey their humble gratitude toward the Director, Maharaja Agrasen Institute of Technology (MAIT), Delhi for giving the required provisions for performing this work.

References

1. Wong, H-S. Philip, Beyond the conventional MOSFET." In *31st European Solid-State Device Research Conference*, pp. 69–72. IEEE, (2001)
2. H.M. Fahad, M.M. Hussain, High-performance silicon nanotube tunneling FET for ultralow-power logic applications. IEEE Trans. Electron Dev. **60**(3), 1034–1039 (2013)
3. C. Hong, J. Zhou, J. Huang, R. Wang, W. Bai, J.B. Kuo, Y. Chen, A general and transformable model platform for emerging multi-gate MOSFETs. IEEE Electron Dev. Lett. **38**(8), 1015–1018 (2017)
4. M. Kumar, S. Haldar, M. Gupta, R.S. Gupta, Impact of gate material engineering (GME) on analog/RF performance of nanowire

Schottky-barrier gate all around (GAA) MOSFET for low power wireless applications: 3D T-CAD simulation. Microelectron. J. **45**(11), 1508–1514 (2014)
5. R.K. Sharma, C.A. Dimitriadis, M. Bucher, A comprehensive analysis of nanoscale single-and multi-gate MOSFETs.". Microelectron. J **52**, 66–72 (2016)
6. Colinge, Jean-Pierre, Abhinav Kranti, Ran Yan, Chi-Woo Lee, Isabelle Ferain, Ran Yu, N. Dehdashti Akhavan, and Pedram Razavi. "Junctionless nanowire transistor (JNT): Properties and design guidelines." *Solid-State Electronics* 65 (2011): 33–37.
7. Y. Pratap, R.S. Subhasis Haldar, Gupta, and Mridula Gupta, , Performance evaluation and reliability issues of junctionless CSG MOSFET for RFIC design. IEEE Trans. Dev. Mater. Reliab. **14**(1), 418–425 (2014)
8. T. Holtij, M. Graef, F.M. Hain, A. Kloes, B. Iñíguez, Compact model for short-channel junctionless accumulation mode double gate MOSFETs. IEEE Trans. Electron Dev. **61**(2), 288–299 (2013)
9. J.H. Choi, T.K. Kim, J.M. Moon, Y.G. Yoon, B.W. Hwang, D.H. Kim, S.-H. Lee, Origin of device performance enhancement of junctionless accumulation-mode (JAM) Bulk FinFETs with high-$\kappa $ Gate Spacers. IEEE Electron Dev. Lett. **35**(12), 1182–1184 (2014)
10. N. Trivedi, M. Kumar, S.S. Subhasis Haldar, M.G. Deswal, R.S. Gupta, Analytical modeling of junctionless accumulation mode cylindrical surrounding gate MOSFET (JAM-CSG). Int. J. Num. Model. Electron. Netw. Dev. Fields **29**(6), 1036–1043 (2016)
11. Zeghbroeck, B.V., Principles of semiconductor devices. (2011). http://ecee.colorado.edu/~bart/book
12. S. Darwin, T.S. Arun, Samuel, A holistic approach on Junctionless dual material double gate (DMDG) MOSFET with high k gate stack for low power digital applications. SILICON **12**(2), 393–403 (2020)
13. S.I. Amin, R.K. Sarin, Charge-plasma based dual-material and gate-stacked architecture of junctionless transistor for enhanced analog performance. Superlatt. Microstruct. **88**, 582–590 (2015)
14. E. Chebaki, F. Djeffal, H. Ferhati, T. Bentrcia, Improved analog/RF performance of double gate junctionless MOSFET using both gate material engineering and drain/source extensions. Superlatt. Microstruct. **92**, 80–91 (2016)
15. A. Sarkar, A.K. Das, S. De, C.K. Sarkar, Effect of gate engineering in double-gate MOSFETs for analog/RF applications. Microelectron. J. **43**(11), 873–882 (2012)
16. Gupta, Sumedha, Neeta Pandey, and R S Gupta, Investigation of dual- material double gate junctionless accumulation-mode cylindrical gate all around (DMDG-JLAM-CGAA) MOSFET with High-k gate stack for low power digital applications. In *2020 IEEE 17th India council international conference (INDICON)*, IEEE (2020)
17. ATLAS, *Device Simulator* (Silvaco International, Santa Clara, 2015)
18. M. Najmzadeh, D. Bouvet, J.-M. Wladyslaw Grabinski, Sallese, and A. M. Ionescu, , Accumulation-mode gate-all-around Si nanowire nMOSFETs with sub-5 nm cross-section and high uniaxial tensile strain. Solid-state Electron. **74**, 114–120 (2012)
19. J. Xiao-Shi, L. Xi, K. Hyuck-In, L. Jong-Ho, A continuous current model of accumulation mode (junctionless) cylindrical surrounding-gate nanowire MOSFETs. Chin. Phys. Lett. **30**(3), 038502 (2013)
20. N.C. Roy, A. Gupta, S. Rai, Analytical surface potential modeling and simulation of junction-less double gate (JLDG) MOSFET for ultra low-power analog/RF circuits. Microelectron. J. **46**(10), 916–922 (2015)
21. T. Holtij, M. Schwarz, A. Kloes, B. Iniguez, Threshold voltage, and 2D potential modeling within short-channel junctionless DG MOSFETs in subthreshold region. Solid-State Electron. **90**, 107–115 (2013)
22. R. Gautam, R.S. Manoj Saxena, Gupta, and Mridula Gupta, , Two dimensional analytical subthreshold model of nanoscale cylindrical surrounding gate MOSFET including impact of localised charges. J. Comput. Theor. Nanosci. **9**(4), 602–610 (2012)
23. C. Jiang, R. Liang, J. Wang, Xu. Jun, A two-dimensional analytical model for short channel junctionless double-gate MOSFETs. AIP Adv. **5**(5), 057122 (2015)
24. C. Li, Ru. Yiqi Zhuang, Han, and Gang Jin, , Subthreshold behavior models for short-channel junctionless tri-material cylindrical surrounding-gate MOSFET. Microelectron. Reliab. **54**(6–7), 1274–1281 (2014)
25. A. Goel, S. Rewari, S. Verma, R.S. Gupta, Physics-based analytic modeling and simulation of gate-induced drain leakage and linearity assessment in dual-metal junctionless accumulation nano-tube FET (DM-JAM-TFET). Appl. Phys. A **126**(5), 1–14 (2020)
26. A. Gnudi, S. Reggiani, E. Gnani, G. Baccarani, Analysis of threshold voltage variability due to random dopant fluctuations in junctionless FETs. IEEE Electron Dev. Lett. **33**(3), 336–338 (2012)
27. J.P. Duarte, S.-J. Choi, Y.-K. Choi, A full-range drain current model for double-gate junctionless transistors.". IEEE Trans. Electron Dev. **58**(12), 4219–4225 (2011)
28. A. Ahmadivand, B. Gerislioglu, Z. Ramezani, Gated graphene island-enabled tunable charge transfer plasmon terahertz metamodulator. Nanoscale **11**(17), 8091–8095 (2019)
29. Trivedi, Nitin, Manoj Kumar, Mridula Gupta, Subhasis Haldar, S. S. Deswal, and R. S. Gupta, Investigation of analog/RF performance of High-k spacer junctionless accumulation-mode cylindrical gate all around (JLAM-CGAA) MOSFET. In *2016 IEEE Uttar Pradesh Section international conference on electrical, computer and electronics engineering (UPCON)*, pp. 201–205. IEEE, (2016)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Genet Resour Crop Evol
https://doi.org/10.1007/s10722-021-01220-6

RESEARCH ARTICLE

Assessment of phytochemical and genetic diversity analysis of *Plumbago zeylanica* L. accessions

Arpita Roy · Neelam Sharma · Navneeta Bharadvaja

Received: 26 September 2020 / Accepted: 22 May 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract Selection of elite accessions is important to get the maximum quantity of bioactive compounds in medicinal plants. In this study, thirteen accessions of *Plumbago zeylanica L.*, a medicinal herb, was studied for phytochemical and genetic diversity analysis. Phytochemical analysis showed that accession number IC-524441 contains the highest amount of flavonoids, phenolic, tannin and plumbagin content, it also possesses maximum antioxidant activity. Genetic diversity analysis showed that 15 SCoT (Start codon targeted) and 20 CBDP (CAAT Box Derived Polymorphism) primers produced 86 and 110 amplicons in total and average were 5.73 and 5.5 amplicons/primer respectively. PIC values ranged from 0.12 to 0.37 and average is 0.25/primer in the case of SCoT marker whereas, in the case of CBDP marker it was 0.15 to 0.59 with average of 0.26/primer. Cluster analysis showed that SCoT marker provides three clusters where cluster 1 and cluster 2 have only 2 accessions while cluster 3 contains nine accessions. Similarly, CBDP marker provides three clusters where cluster 1 contains only 1 accession and cluster 2 and cluster 3 contains six accessions in each. Further AMOVA analysis of SCoT marker possesses maximum variation of 8% among agro-ecological regions whereas in case of CBDP marker its 5%. From the present investigation, it was found that IC-524441 is an elite accession and contains highest amount of bioactive compound. Our results also demonstrate that both SCoT and CBDP markers are informative and can be utilized for evaluation of genetic relationships among *Plumbago zeylanica* L. accessions.

Keywords *Plumbago zeylanica* L. · Phytochemicals · Genetic diversity · SCoT · CBDP

A. Roy (✉) · N. Bharadvaja (✉)
Department of Biotechnology, Delhi Technological University, New Delhi, India
e-mail: arbt2014@gmail.com

N. Bharadvaja
e-mail: navneetab@dce.ac.in

A. Roy
Department of Biotechnology, School of Engineering & Technology, Sharda University, Greater Noida, India

N. Sharma
Tissue Culture and Cryopreservation Unit, National Bureau of Plant Genetic Resources, New Delhi, India

Introduction

Plumbago zeylanica L. (*Plumbaginaceae*) is a valuable medicinal plant commonly known as Chitrak in Hindi and White leadwort in English (Roy and Bharadvaja 2019). It has been extensively utilized for the treatment of secondary syphilis, anasarca, diarrhoea, rheumatism, leprosy, and piles (Roy and Bharadvaja 2017; Roy and Bharadvaja 2018). Diverse

Published online: 02 June 2021

therapeutic activity of this plant is due to phytocompounds such as alkaloids, flavonoids, steroids, sterols, naphthoquinones, etc (Roy and Bharadvaja 2019). It possesses various pharmaceutical applications such as neuroprotective, antibacterial, antifungal, antimalarial, antiviral, anti-cancer, anti-plasmodial, hepatoprotective, anti-inflammatory, anti-hyperglycemic, anti-atherosclerotic activity (Roy and Bharadvaja 2017a, b; Jain et al. 2020; Jayanthi et al. 2020; Shukla et al. 2020). Increase demand for natural product in domestic and international markets validates the requirement of phytochemical analysis in *P. zeylanica*. Production of bioactive compounds are influenced by environmental factors; therefore, studying different accessions for phytochemical analysis provides the details about best accession. Best accession can be further utilized by the pharmaceutical industry for enhancing the production of quality drugs. Various studies reported phytochemical analysis in *P. zeylanica* (Saxena et al. 2016; Jain et al. 2020). However, none of them utilized different accessions to estimate the presence of different phytochemical contents.

The continuous exploitation and growing demand of *P. zeylanica* have given rise to concerns like sustainable use, conservation and authenticity of the material. Non-availability of consolidated and reliable information about the genetic variation of the plant in commercial trade accentuated the concern. Thus, the correct knowledge of genetic diversity of *P. zeylanica* is necessary. For long-term utilization of *P. zeylanica*, it is necessary to understand the extent of inherited variability and intraspecific similarity. Genetic diversity may present in the plants as a result of evolutionary process. However, environmental conditions play an important role in creating genetic diversity. Climatic conditions vary in different parts of India and thus India is divided into 12 agro-climatic zones. Plants belong to different agro-climatic zones differ in genetic make-up. Understanding of genetic diversity and reproduction behaviour of plants is necessary for initiating any conservation and management program, which is important for preserving genetic diversity within the population in the plants like *P. zeylanica*. Genetic variability is responsible for variation among different accessions and helps in the management of current resources to reveal genetic associations among various accessions.

Molecular genetic markers represent an effective technique to survey the effect of different parameters on genetic diversity (Nadeem et al. 2018; Kumar and Agrawal 2019). In molecular genetic studies, DNA markers have been used for the improvement of crops. Genomic sequence has huge potential for development of new markers in different plant species. But selection of markers mainly depends on plant species, technical knowledge, and research facilities of working laboratories (Abdin et al. 2017). Several markers have been utilized for the genetic diversity analysis such as AFLP, ISSR, RAPD, SSR, SCoT, CBDP, etc. (Gorji et al. 2011; Tiwari et al. 2016; Verma et al. 2017; Kumar and Agrawal 2019). Among them, SCoT and CBDP have been proved to be helpful for population genetic studies. SCoT primer is single 18-mer oligonucleotides which was used to investigate conserved area of gene which comprises of translation initiation codon i.e., ATG (Gupta et al. 2019). SCoT marker examination has been effectively utilized in various genetic diversity studies such as *Elymus sibiricus, Andrographis paniculate,* rose varieties, etc. (Zhang et al. 2015; Tiwari et al. 2016; Agarwal et al. 2019). CBDP is an 18 nucleotides base single primer, comprise of a central CCAAT nucleotides core which is flanking with the filler sequence at 5′ end and di- or trinucleotides at 3′ end (Singh et al. 2014). Genetic diversity in *Plumbago zeylanica* L. using different markers such as SSR, ITS, ISSR, RAPD, etc. has been reported in the literature (Haji et al. 2014; Panda et al. 2015). In the present study phytochemical analysis and genetic diversity analysis of thirteen accessions of *P. zeylanica* has been performed to select best accession.

Material and method

Collection of material

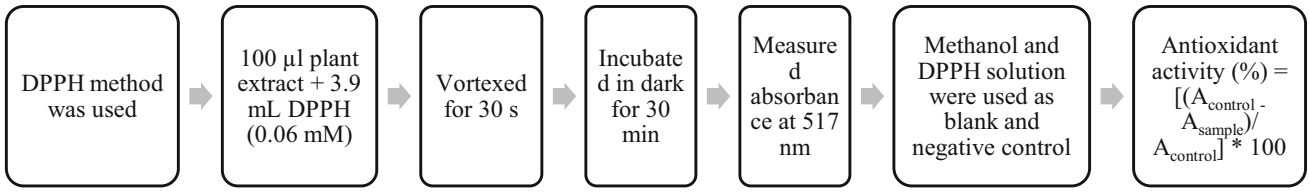
13 accessions of *P. zeylanica* were collected from National Bureau of Plant Genetic Research (NBPGR) and were maintained under *in-vitro* conditions.

Biochemical analysis

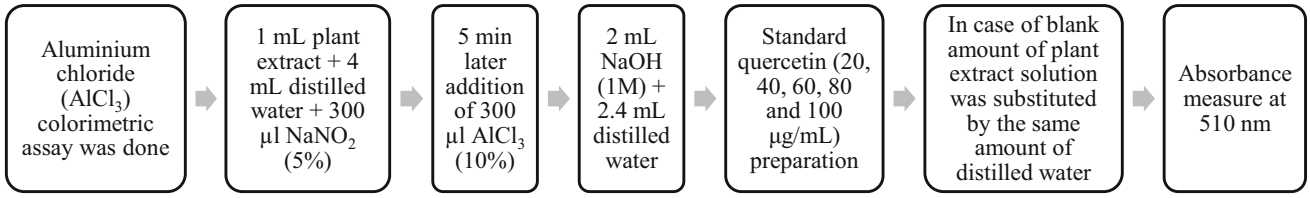
For estimation of plumbagin, phenol, tannin, flavonoid content and antioxidant activity (Roy and Bharadvaja 2018) in vitro cultures of different accessions of *P. zeylanica* were used. Freshly *in-vitro* grown plant material was dried and grounded. Further 10.0 gm of

sample was soaked in 10 ml methanol (concentration 1gm/ml) for 48 h and sample was filtered.

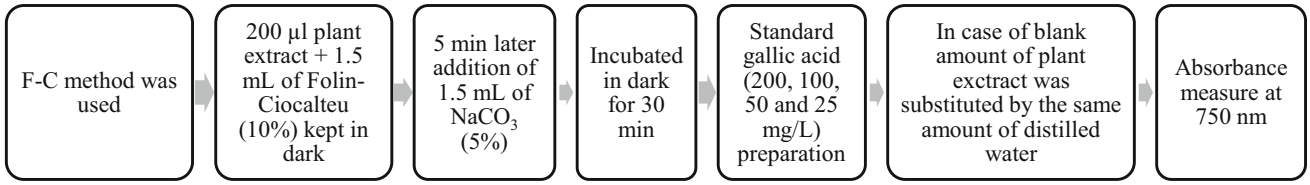
Antioxidant activity



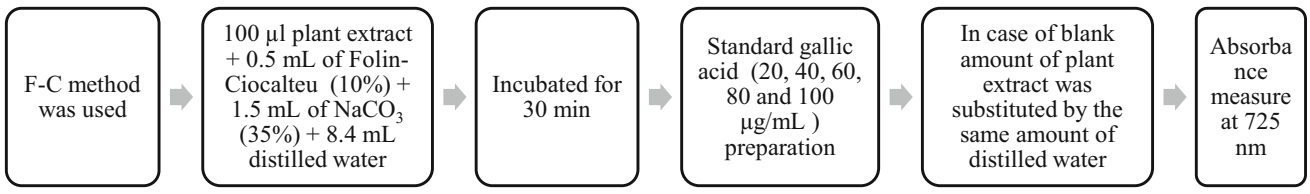
Total flavonoid estimation



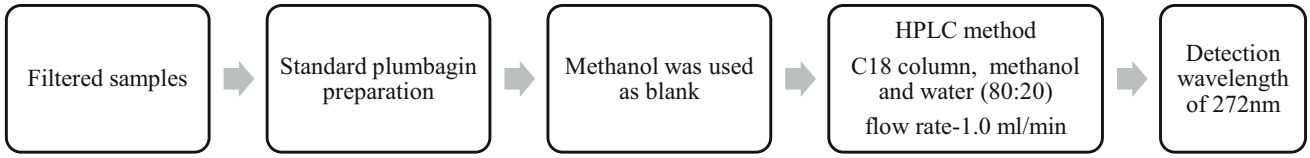
Total phenolic estimation



Total tannin estimation



Plumbagin estimation



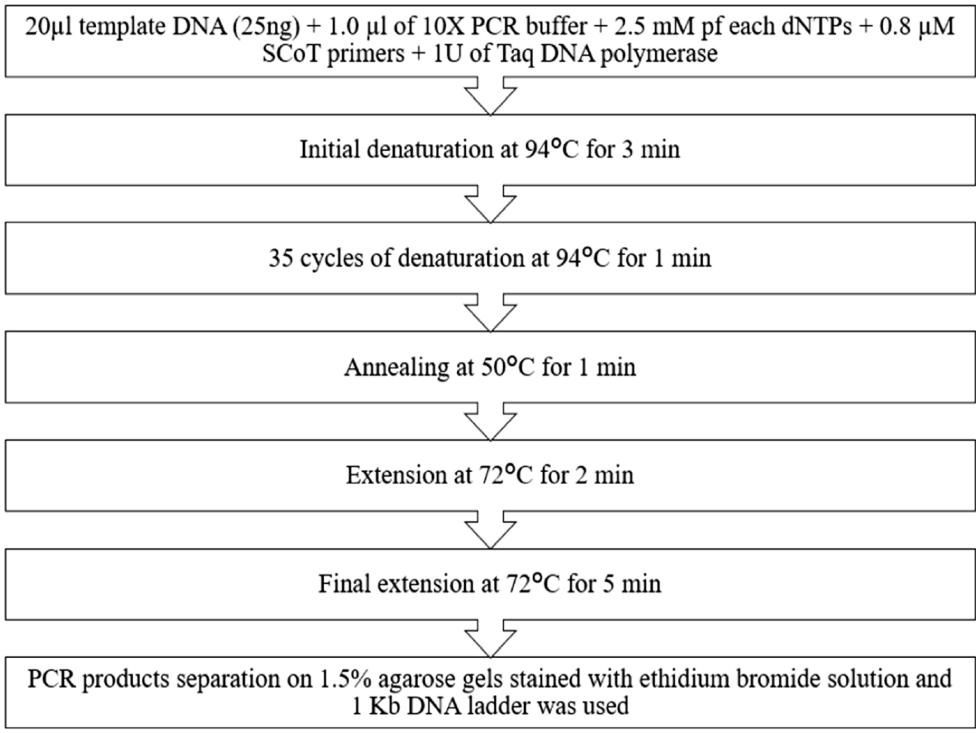
Genetic diversity analysis

Extraction of DNA

Fresh leaves were used for genomic DNA extraction. CTAB method was used for extraction of DNA. Approximately 200 mg of leaves were ground with liquid nitrogen and mixed with 2 ml of CTAB (2% pre-warmed CTAB containing 0.2% β-mercaptoethanol). Then samples were poured into eppendorf tubes and incubated at 68 °C for 45 min in water bath. 1 ml of chloroform: isoamyl alcohol (24:1) was added before centrifugation at 10,000 rpm for10 mins. Supernatant was transferred to a fresh tube. Then an equal volume of ice-cold isopropanol was added and samples were kept at 4°C overnight. Samples were centrifuged (Eppendorf Centrifuge 5804 R) at 10,000 rpm for 10 min to form a pellet. Pellet was washed with 70% ethanol and air dried, finally dissolved in 100 µl TE. RNAse treatment was given to the samples and final DNA samples were stored at -20 °C. DNA quality was verified using spectrophotometry (Perkin Elmer UV vis spectrophotometer LAMBDA™ 365) at 260 nm and 280 nm and electrophoresis (0.8% agarose gel).

SCoT and CBDP markers analysis

Total 15 SCoT and 20 CBDP primers (Sigma Chemicals, USA) were used to study polymorphism in *P. zeylanica* and 1 Kb DNA ladder (Fermentas, USA) was used.


20µl template DNA (25ng) + 1.0 µl of 10X PCR buffer + 2.5 mM pf each dNTPs + 0.8 µM SCoT primers + 1U of Taq DNA polymerase
Initial denaturation at 94°C for 3 min
35 cycles of denaturation at 94°C for 1 min
Annealing at 50°C for 1 min
Extension at 72°C for 2 min
Final extension at 72°C for 5 min
PCR products separation on 1.5% agarose gels stained with ethidium bromide solution and 1 Kb DNA ladder was used

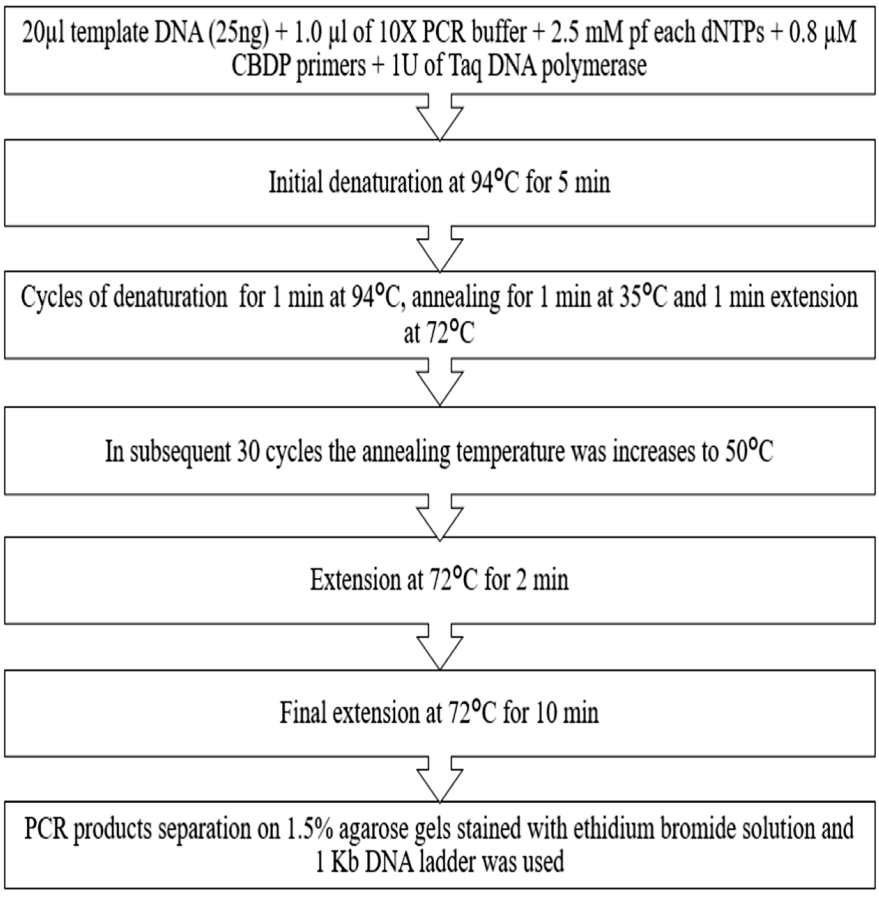

20µl template DNA (25ng) + 1.0 µl of 10X PCR buffer + 2.5 mM pf each dNTPs + 0.8 µM CBDP primers + 1U of Taq DNA polymerase
Initial denaturation at 94°C for 5 min
Cycles of denaturation for 1 min at 94°C, annealing for 1 min at 35°C and 1 min extension at 72°C
In subsequent 30 cycles the annealing temperature was increases to 50°C
Extension at 72°C for 2 min
Final extension at 72°C for 10 min
PCR products separation on 1.5% agarose gels stained with ethidium bromide solution and 1 Kb DNA ladder was used

Table 1 Biochemical analysis of 13 accessions of *P. zeylanica*

Accession no	Antioxidant activity (μg/ml)	Total phenolic content estimation(μg/ml)	Total flavonoid estimation (μg/ml)	Total tannin estimation (μg/ml)	Plumbagin estimation (mg/g)
IC-524440	89.20 ± 0.72^{c}	69.23 ± 0.78^{c}	55.70 ± 0.96^{b}	94.53 ± 0.50^{b}	22.44 ± 0.71^{b}
IC-524441	134.17 ± 0.76^{a}	80.47 ± 0.64^{a}	78.63 ± 0.78^{a}	109.50 ± 0.50^{a}	25.54 ± 0.58^{a}
IC-524444	109.50 ± 0.87^{b}	76.97 ± 0.55^{b}	62.17 ± 0.86^{a}	94.97 ± 0.85^{b}	24.17 ± 0.58^{ab}
IC-539866	19.13 ± 0.81 ^{h}	40.13 ± 0.71^{e}	18.40 ± 0.78^{f}	50.60 ± 0.60^{f}	7.79 ± 0.64^{i}
IC-539867	15.67 ± 0.58 ^{h}	30.50 ± 0.70^{f}	30.47 ± 0.55^{d}	73.07 ± 0.70^{d}	5.3 ± 0.55^{j}
IC-256070	29.50 ± 0.50 ^{g}	61.57 ± 0.55^{c}	50.33 ± 0.76^{b}	89.00 ± 0.50^{c}	7.98 ± 0.61^{i}
IC-211317	78.33 ± 0.58^{d}	45.00 ± 0.50^{e}	51.30 ± 0.70^{b}	51.47 ± 0.55	16.89 ± 0.55^{d}
IC-421418	79.57 ± 0.75^{d}	55.13 ± 0.71^{d}	43.87 ± 0.78^{c}	86.00 ± 0.40^{c}	21.18 ± 0.76^{c}
IC-340668	54.67 ± 0.58^{f}	34.97 ± 0.35^{f}	25.13 ± 0.61^{e}	66.93 ± 0.40^{e}	8.74 ± 0.96 ^{h}
IC-0624282	74.30 ± 0.61^{d}	66.67 ± 0.49^{c}	29.03 ± 0.55^{e}	55.20 ± 0.62^{f}	10.74 ± 0.55^{f}
IC-0624281	65.17 ± 0.76^{e}	70.80 ± 0.98^{b}	49.33 ± 0.67^{c}	78.07 ± 0.80^{d}	10.07 ± 0.64^{f}
IC-398891	59.63 ± 0.64^{f}	53.03 ± 0.55^{d}	47.17 ± 0.38^{c}	54.97 ± 0.85^{f}	9.64 ± 0.71 ^{g}
IC-439214	64.87 ± 0.81^{e}	49.00 ± 0.50^{e}	31.30 ± 0.62^{d}	60.70 ± 0.62^{e}	13.44 ± 0.58^{e}

Different letters (a–j) indicate significant differences ($P < 0.05$) within groups using one-way ANOVA

Table 2 SCoT primers sequence analysis

Primer	Primer sequence	% of GC	Total no of amplicon	No. of monomorphic amplicon	No. of polymorphic amplicon	Polymorphism %	PIC
SCoT1	ACGACATGGCGACCACGC	66.67	4	1	3	75	0.27
SCoT2	ACCATGGCTACCACCGAC	61.11	4	1	3	75	0.27
SCoT3	ACCATGGCTACCACCGAG	61.11	4	1	3	75	0.31
SCoT4	ACCATGGCTACCACCGGG	66.67	3	2	1	33.33	0.12
SCoT5	ACCATGGCTACCACCGTG	61.11	7	1	6	85.71	0.32
SCoT6	CCATGGCTACCACCGGCC	72.22	7	2	5	71.43	0.23
SCoT7	CCATGGCTACCACCGCAG	66.67	7	3	4	57.14	0.23
SCoT8	CAACAATGGCTACCACCA	50	6	1	5	83.33	0.29
SCoT9	CAACAATGGCTACCACCC	55.56	4	0	4	100.00	0.28
SCoT10	CAACAATGGCTACCACCG	55.56	6	1	5	83.33	0.27
SCoT11	CAACAATGGCTACCACCT	50	5	2	3	60	0.22
SCoT12	CAACAATGGCTACCACGT	50	5	1	4	80	0.18
SCoT13	CAACAATGGCTACCAGCA	50	7	0	7	100	0.19
SCoT14	ACGACATGGCGACCAACG	61.11	7	1	6	85.71	0.37
SCoT15	CCATGGCTACCACCGGCG	77.78	10	2	8	80	0.22

Data analysis The product of PCR (Eppendorf Mastercycler Nexus Thermal Cyclers) analysis was scored visually. Only clear noticeable bands were taken into consideration and presence of band was scored as "1" whereas absence of band as "0" for binary data matrix. Only reproducible bands were used for analysis. Discriminatory power of the primers was evaluated using Polymorphic information content

(PIC) where PIC value was calculated as PIC = 2*fi*(1 − fi) where fi is amplified allele frequency and (1 − fi) is non-amplified allele frequency. Cluster analysis using UPGMA algorithm and AMOVA was performed.

Results and discussion

Biochemical analysis

Antioxidant activity was performed using DPPH method. DPPH free radical scavenging effect of different accession of *P. zeylanica* ranges from 15.67–134 µg/ml (Table 1). All the extracts have a notable effect on scavenging free radicals and IC-524441 showed highest scavenging activity (134.17 ± 0.76 µg/ml) while lowest scavenging activity was observed in IC-539867 (15.67 ± 0.58 µg/ml). This could be due to highest content of total polyphenolic compounds which include anthraquinones, flavonoids, xanthones, anthocyanidins and tannins (Panche et al. 2016).

Total phenolic content of samples ranges from 30.50 to 80.47 µg/ml (Table 1). IC-524441 showed highest amount of phenolic content (80.47 ± 0.64 µg/ml) whereas lowest content was found in IC-539867 (30.50 ± 0.70 µg/ml). Presence of different phenolic compounds may be the cause of variation of total phenolic content in the plant extracts. Polyphenols are vital dietary antioxidants as they possess strong free radical scavenging activity. Various reports suggested that polyphenols protect against numerous diseases (Cory et al. 2018). These compounds show a significant role in growth and reproduction, and protect against damaging pathogens. Sharma et al. (2014) reported that methanolic extract of *P. zeylanica* leaves also showed presence of phenolic content i.e., 28.25 ± 0.001 µg/ml.

Total flavonoid content ranges from 18.40–68.63 µg/ml (Table 1). IC-524441 exhibits maximum amount of flavonoid content

Table 3 CBDP primers sequence analysis

Primer	Primer sequence	% of GC	Total no of amplicon	No. of monomorphic amplicon	No. of polymorphic amplicon	Polymorphism %	PIC
CAAT1	TGAGCACGATCCAATACC	50.00	5	2	3	60.0	0.18
CAAT2	TGAGCACGATCCAATAAG	44.44	9	2	7	77.8	0.59
CAAT3	TGAGCACGATCCAATCAG	50.00	5	1	4	80.0	0.16
CAAT4	TGAGCACGATCCAATGAT	44.44	6	0	6	100.0	0.26
CAAT5	TGAGCACGATCCAATTGC	50.00	3	1	2	66.7	0.21
CAAT6	TGAGCACGATCCAATGAG	50.00	4	1	3	75.0	0.31
CAAT7	TGAGCACGATCCAATGCG	55.56	8	0	8	100.0	0.32
CAAT8	TGAGCACGATCCAATTGA	44.44	4	1	3	75	0.35
CAAT9	TGAGCACGATCCAATTCA	44.44	5	0	5	100	0.36
CAAT10	CTGAGCACGATCCAATAG	50.00	5	1	4	80	0.23
CAAT11	CTGAGCACGATCCAATAC	50.00	8	2	6	75	0.22
CAAT12	CTGAGCACGATCCAATAT	44.44	5	0	5	100	0.27
CAAT13	CTGAGCACGATCCAATCA	50.00	7	1	6	85.71	0.24
CAAT14	CTGAGCACGATCCAATGG	55.56	4	1	3	75	0.17
CAAT15	CTGAGCACGATCCAATGT	50.00	4	1	3	75	0.25
CAAT16	TGAGCACGATCCAATAGC	50.00	4	0	4	100	0.23
CAAT17	TGAGCACGATCCAATAAT	38.89	4	0	4	100	0.23
CAAT18	TGAGCACGATCCAATGTT	44.44	5	0	5	100	0.33
CAAT19	TGAGCACGATCCAATATA	44.44	8	1	7	87.5	0.15
CAAT20	CTGAGCACGATCCAATAG	50.00	7	1	6	85.71	0.25

(78.63 ± 0.78) while minimum content was found in IC-539866 (18.40 ± 0.78 µg/ml). Flavonoids are potential scavengers of most oxidizing molecules which include singlet oxygen and several other free radicals involved in various diseases such as cancer (Kumar and Pandey 2013). It suppresses reactive oxygen formation, chelating trace elements and protects antioxidant defenses (Kurutas 2015). Sharma et al. (2014) reported flavonoid content in stem of *P. zeylanica* i.e., 72.3 µg/ml.

Total tannin content was determined using gallic acid as standard and varied between 50.60–109.50 µg/ml (Table 1). Highest amount was observed in accession number IC-524441 i.e., 109.50 ± 0.50 µg/ml and lowest in accession number 539866 which is 50.60 ± 0.60 µg/ml. Tannins are water-soluble compounds, widely found in nature and possibly found in all plants (Pizzi 2019). Tannins possess wide range of pharmaceutical activities such as antimicrobial, antioxidant, free radical scavenging and anti-ulcerogenic activities (Shad et al. 2012).

The amount of plumbagin was estimated by comparing samples with standard plumbagin solution. Plumbagin concentration varied between 5.3 and 25.54 µg/ml (Table 1). Results showed that accession numbers IC-524441 contained maximum amount of plumbagin i.e., 25.54 ± 0.58 µg/ml and minimum in accession number 539867 which is 5.3 ± 0.55 µg/ml. Accession number IC-524441 showed maximum phenolic content, flavonoid content, tannin content, antioxidant activity and plumbagin production.

Genetic diversity analysis

To evaluate genetic diversity among 13 accessions of *P. zeylanica*, two different molecular markers i.e., CBDP and SCoT were used. SCoT and CBDP are gene targeted markers. This is the first report on gene-targeted markers that were used for genetic diversity analysis in different *P. zeylanica* accessions.

Fifteen SCoT primers were used for DNA amplification. All 15 primers generated polymorphic DNA amplification products (Table 2). These fifteen SCoT primers produce 86 amplicons and number of amplicons ranged from 3(SCoT-4) to 10(SCoT-15) and average is 5.73 amplicons/primer. Polymorphic bands vary from 1 to 8 and average 4.45 amplicons/primer. Polymorphism percentage varies from 33.33 to 100% with an average of 76.33%/primer. PIC values vary from 0.12 to 0.37 and average is 0.25/primer. Similarly, in case of *Tinospora cardifolia*, 102 amplicons (5.3 amplicons/primer) were obtained with nineteen SCoT primers (Paliwal et al. 2013). They reported that average polymorphic amplicons were 4.74/ primer and average polymorphism and PIC value were 87.02% and 0.49/ primer. Satya et al. (2015) observed that 20 SCoT primers generated 136 amplicons (6.8 amplicons/primer) with 87.5% polymorphism in *Boehmeria*

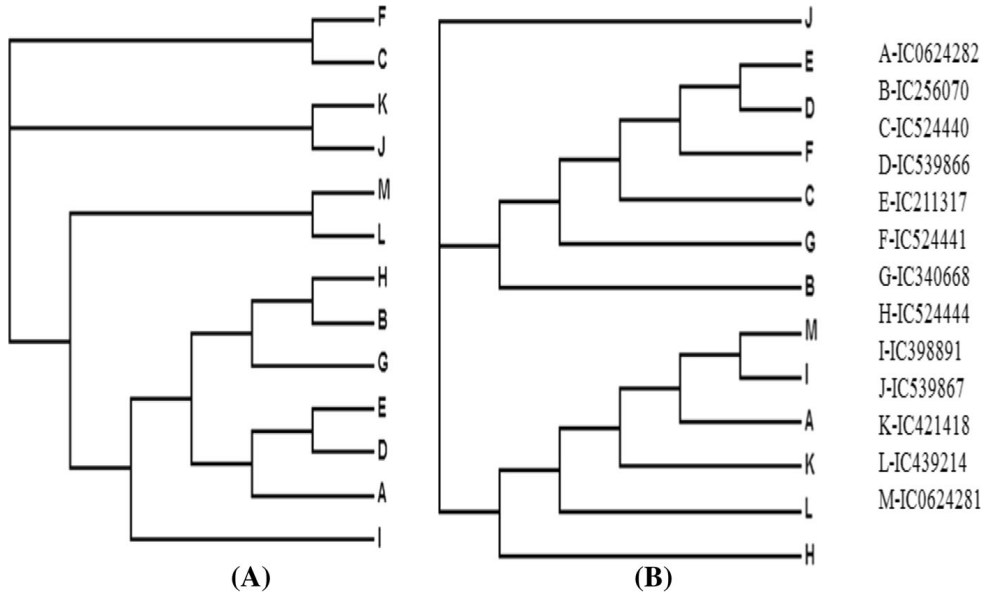


Fig. 1 UPGMA clustering of 13 accessions of *P. zeylanica* for SCoT and CBDP markers data

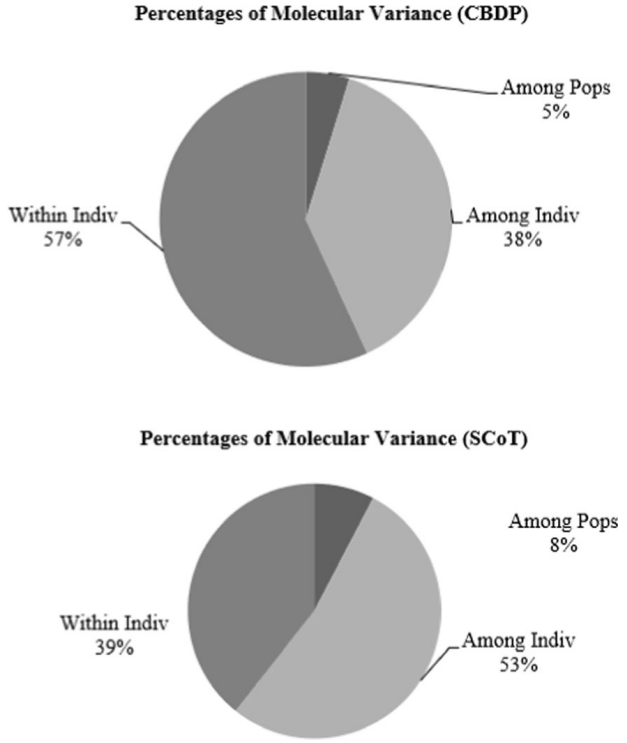


Fig. 2 Analysis of Molecular Variance (AMOVA) with CBDP and SCoT marker

nivea L. Gaudich genotypes. Tiwari et al. (2016) reported that in case of *Andrographis paniculata,* 20 SCoT primers generate 132 amplicons (6 amplicons/ primer) and average polymorphic amplicons of 4.5/ primer. The percentage of polymorphism was 75% and average PIC value was 0.34/ primer. These outcomes indicated that SCoT primers showed high amplification potential and polymorphism. Gholamian et al. (2019) used 19 SCoT primers in *Triticum urartu* Tumanian ex Gandilyan and 19 SCoT primers generated 185 polymorphic bands with an average of 9.74/ primer, and PIC ranged from 0.41 to 0.50. Vanijajiva (2020) reported that 27 SCoT primers generate 176 amplicons and 101 (57.38%) of them were polymorphic in case of *Manilkara* accessions. Based on similarity coefficients, total of 36 accessions were clustered into two groups.

Twenty CBDP primers generated 110 amplicons (5.5 amplicons/primer (Table 3). Polymorphic amplicons vary from 2(CAAT-5) to 8 (CAAT-7) and average is 4.7 amplicons/ primer. Polymorphism percentage vary from 60.00 to 100% with average of 84.9%/primer. PIC values vary from 0.15 (CAAT-19) to 0.59 (CAAT-2), average is 0.26/primer. Similarly, using seventeen CBDP primers, a study reported gender specific diversity in jojoba genotypes where 133 amplicons in male and 126 amplicons in female (Heikrujam et al. 2015). They reported that in female Jojoba average PIC was 0.42/ CBDP primer. Singh et al. (2014) used twenty-five CBDP markers in Jute and obtained 82% polymorphism and average PIC value of 0.41/primer. In another study where Chickpea genotypes were used, the average PIC value was 0.45/ primer with CBDP markers (Hajibarat et al. 2015). Etminan et al. (2019) performed genetic diversity analysis of *Aegilops* and *Triticum* species using CBDP marker and 15 CBDP primers generated 141 polymorphic bands with an average of 9.40/primer. Gholamian et al. (2019) used 15 CBDP markers in *Triticum urartu* Tumanian ex Gandilyan and 15 CBDP primers generated 141 polymorphic bands with an average of 9.40/primer, and PIC ranged from 0.40 to 0.49.

Cluster analysis

Cluster analysis was done to estimate the genetic diversity among *P. zeylanica* accessions. In case of SCoT data analysis, three clusters were obtained in dendrogram. In cluster 1 and cluster 2 only 2 accessions were clustered while cluster 3 contains nine accessions. Cluster 3 was further grouped into two subclusters. Paliwal et al. (2013) also found three clusters using SCoT data in case of *T. cardifolia*. In cluster1 and 3 three and four accessions were present while in cluster 2, thirty-two accessions were grouped (Fig. 1). In case of CBDP data analysis, three clusters were obtained in dendrogram. In cluster 1, only 1 accession was grouped while cluster 2 and cluster 3, six accessions were present in each cluster. Cluster 2 was further grouped into five subclusters where in subcluster 1, 2 accessions were present and other four subclusters contain 1 accession in each. Similarly, Cluster 3 was further grouped into five subclusters where in subcluster 1, 2 accessions were present and other four subcluster contains 1 accession in each. Heikrujam et al. (2015) reported that in Jojoba genotypes, clusters were grouped into two major groups, one group consisting of all female genotypes and another one consisting of all male genotypes. Gholamian et al. (2019) reported that in *Triticum urartu* Tumanian ex Gandilyan 72 accessions were grouped into two main clusters based on the SCoT and CBDP markers. Etminan et al. (2019) reported that in

Aegilops and *Triticum* species accessions within each cluster were separated into distinct groups where 228 accessions were separated into six clusters. Vanijajiva (2020) reported that a total of 36 accessions were clustered into two groups in *Manilkara* accessions.

AMOVA study

AMOVA was done to estimate the difference in *P. zeylanica* accessions. AMOVA analysis of SCoT marker showed variation of 53% among individuals, 39% within individuals and 8% among agro-ecological regions. Similarly, AMOVA analysis of CBDP marker showed maximum variation of 57% within individuals, 38% among individuals and 5% among agro ecological regions (Fig. 2). SCoT and CBDP markers are able to provide information about DNA polymorphism. A study reported high level of genetic diversity in case of *Breonandia salicinia* population i.e., 17% with 14 ISSR markers (Gaafar et al. 2014). Gorji et al. (2011) studied twenty-four varieties of tetraploid potato using SCoT, RAPD and ISSR markers and performed AMOVA analysis, they observed less than 10% of total genetic variation. Tiwari et al. (2016) reported that AMOVA analysis for CBDP marker showed maximum variation of 10% whereas in case SCoT markers it was 2%.

Conclusion

Production of bioactive compounds is influenced by various factors. Therefore, studying different accessions for biochemical analysis provides the details about best accession. Among the studied accessions, accession number IC-524441 provides maximum antioxidant activity, flavonoids content, phenolic content, tannin content and plumbagin content. SCoT and CBDP marker analysis were effectively done to evaluate genetic diversity among *P. zeylanica* accessions. Polymorphism shown by SCoT and CBDP can be utilized for molecular study of *P. zeylanica* accessions, which provides valuable information for germplasm management, improve breeding strategies, and genetic resources conservation. Both the markers provide higher reproducibility and can be used for any plant species. Therefore, based on the requisite and plant species, several numbers of primers can be designed. This method has numerous advantages and it is comparatively easy to carry out as there is no requirement of previous genome sequence information.

Acknowledgements We would like to thank our department for providing required facilities and Director NBPGR (NBPGR) for providing the plant material. We would also like to thank Dr. Rakesh Singh(NBPGR) for helping in AMOVA analysis.

Funding Funding was provided by Delhi Technological University.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Abdin MZ, Arya L, Verma M (2017) Use of SCoT markers to assess the gene flow and population structure among two different populations of bottle gourd. Plant Gene 9:80–86

Agarwal A, Gupta V, Haq SU, Jatav PK, Kothari SL, Kachhwaha S (2019) Assessment of genetic diversity in 29 rose germplasms using SCoT marker. J King Saud Univ Sci 31(4):780–788

Cory H, Passarelli S, Szeto J, Tamez M, Mattei J (2018) The role of polyphenols in human health and food systems: a mini-review. Front Nutr 5:87

Etminan A, Pour-Aboughadareh A, Mehrabi AA, Shooshtari L, Ahmadi-Rad A, Moradkhani H (2019) Molecular characterization of the wild relatives of wheat using CAAT-box derived polymorphism. Plant Biosyst Int J Dealing All Aspects Plant Biol 153(3):398–405

Gaafar ARZ, Al-Qurainy F, Khan S (2014) Assessment of genetic diversity in the endangered populations of *Breonadia salicina* (Rubiaceae) growing in The Kingdom of Saudi Arabia using inter-simple sequence repeat markers. BMC Genet 15(1):109

Gholamian F, Etminan A, Changizi M, Khaghani S, Gomarian M (2019) Assessment of genetic diversity in Triticum urartu Thumanjan ex Gandilyan accessions using start codon targeted polymorphism (SCoT) and CAAT-box derived polymorphism (CBDP) markers. Biotechnol Biotechnol Equip 33(1):1653–1662

Gorji AM, Poczai P, Polgar Z, Taller J (2011) Efficiency of arbitrarily amplified dominant markers (SCoT, ISSR and RAPD) for diagnostic fingerprinting in tetraploid potato. Am J Potato Res 88(3):226–237

Gupta V, Jatav PK, Haq SU, Verma KS, Kaul VK, Kothari SL, Kachhwaha S (2019) Translation initiation codon (ATG) or SCoT markers-based polymorphism study within and across various Capsicum accessions: insight from their amplification, cross-transferability and genetic diversity. J Genet 98(2):61

Haji RFA, Bhargava M, Akhoon BA, Kumar A, Brindavanam NB, Verma V (2014) Correlation and functional differentiation between different markers to study the genetic

diversity analysis in medicinally important plant *Plumbago zeylanica*. Ind Crops Prod 55:75–82
Hajibarat Z, Saidi A, Hajibarat Z, Talebi R (2015) Characterization of genetic diversity in chickpea using SSR markers, start codon targeted polymorphism (SCoT) and conserved DNA-derived polymorphism (CDDP). Physiology and molecular biology of plants 21(3):365–373
Heikrujam M, Kumar J, Agrawal V (2015) Genetic diversity analysis among male and female Jojoba genotypes employing gene targeted molecular markers, start codon targeted (SCoT) polymorphism and CAAT box-derived polymorphism (CBDP) markers. Meta Gene 5:90–97
Jain P, Sharma HP, Singh P (2020) Anti-fungal, anti-oxidant and phytochemical analysis of *Plumbago zeylanica* L. Vegetos 1–11
Jayanthi M, Gokulanathan A, Haribalan P, Ashakiran K, Kumar CD, Kamla D, Renukadevi BS (2020) Plumbagin from two Plumbago species inhibits the growth of stomach and breast cancer cell lines. Ind Crops Prod 146:112147
Kumar J, Agrawal V (2019) Assessment of genetic diversity, population structure and sex identification in dioecious crop, *Trichosanthes dioica* employing ISSR. SCoT SRAP Markers Heliyon 5(3):e01346
Kumar S, Pandey AK (2013) Chemistry and biological activities of flavonoids: an overview. Sci World J 2013:162750
Kurutas EB (2015) The importance of antioxidants which play the role in cellular response against oxidative/nitrosative stress: current state. Nutr J 15(1):71
Nadeem MA, Nawaz MA, Shahid MQ, Doğan Y, Comertpay G, Yıldız M, Özkan H (2018) DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. Biotechnol Biotechnol Equip 32(2):261–285
Paliwal R, Singh R, Singh AK, Kumar S, Kumar A, Majumdar RS (2013) Molecular characterization of Giloe (*Tinospora cordifolia* Willd. Miers ex Hook. F. and Thoms.) accessions using start codon targeted (SCoT) markers. Int J Med Arom Plants 3(4):413–422
Panche AN, Diwan AD, Chandra SR (2016) Flavonoids: an overview. J Nutr Sci 5:e47
Panda S, Naik D, Kamble A (2015) Population structure and genetic diversity of the perennial medicinal shrub *Plumbago*. AoB Plants 7:plv048
Pizzi A (2019) Tannins: prospectives and actual industrial applications. Biomolecules 9(8):344
Roy A, Bharadvaja N (2017a) A review on pharmaceutically important medical plant: *Plumbago zeylanica*. J Ayurved Herbal Med 3(4):225–228
Roy A, Bharadvaja N (2017b) Silver nanoparticles synthesis from a pharmaceutically important medicinal plant *Plumbago zeylanica*. MOJ Bioequiv Availab 3(5):00046
Roy A, Bharadvaja N (2018) Effect of various culture conditions on shoot multiplication and GC–MS analysis of *Plumbago zeylanica* accessions for plumbagin production. Acta Physiol Plant 40(11):190
Roy A, Bharadvaja N (2019) Silver nanoparticle synthesis from Plumbago zeylanica and its dye degradation activity. Bioinspired, Biomimetic and Nanobiomaterials 8(2):130–140
Roy A, Bharadvaja N (2017) Establishment of the shoot and callus culture of an important medicinal plant Plumbago zeylanica. Adv Plants Agric Res 7(5):00274
Roy A, Bharadvaja N (2019) Establishment of root suspension culture of *Plumbago zeylanica* and enhanced production of plumbagin. Ind Crops Prod 137:419–427
Satya P, Karan M, Jana S, Mitra S, Sharma A, Karmakar PG, Ray DP (2015) Start codon targeted (SCoT) polymorphism reveals genetic diversity in wild and domesticated populations of ramie (*Boehmeria nivea* L. Gaudich.), a premium textile fiber producing species. Meta Gene 3:62–70
Saxena A, Gautam S, Arya KR, Singh RK (2016) Comparative study of phytochemicals, antioxidative potential & activity of enzymatic antioxidants of *Eclipta alba* and *Plumbago zeylanica* by in vitro Assays. Free Radicals Antioxid 6(2)
Shad MA, Haq N, Tanzila R, Ahmad HB, Mazhar H (2012) Optimization of extraction efficiency of tannins from *Cichorium intybus* L.: application of response surface methodology. J Med Plants Res 6(28):4467–4474
Sharma I, Mathur M, Singh GP, Rishi A (2014) Quantitative estimation of phenolic and flavonoid content and antioxidant activity of various extracts of different parts of *Plumbago zeylanica Linn*. Int J Drug Dev Res 6(2):136–140
Shukla PK, Misra A, Patra KK, Srivastava S (2020) Study of metabolite variability in *Plumbago zeylanica* Linn. collected from different localities of the Gangetic plains of India. JPC J Planar Chromatogr Modern TLC 33:179–189
Singh AK, Rana MK, Singh S, Kumar S, Kumar R, Singh R (2014) CAAT box-derived polymorphism (CBDP): a novel promoter-targeted molecular marker for plants. J Plant Biochem Biotechnol 23(2):175–183
Tiwari G, Singh R, Singh N, Choudhury DR, Paliwal R, Kumar A, Gupta V (2016) Study of arbitrarily amplified (RAPD and ISSR) and gene targeted (SCoT and CBDP) markers for genetic diversity and population structure in Kalmegh [*Andrographis paniculata* (Burm. f.) Nees]. Ind Crops Prod 86:1–11
Vanijajiva O (2020) Start codon targeted (SCoT) polymorphism reveals genetic diversity of Manilkara in Thailand. Biodiversitas 21:666–673
Verma KS, Ul Haq S, Kachhwaha S, Kothari SL (2017) RAPD and ISSR marker assessment of genetic diversity in *Citrullus colocynthis* (L.) Schrad: a unique source of germplasm highly adapted to drought and high-temperature stress. 3 Biotech 7(5):288
Zhang J, Xie W, Wang Y, Zhao X (2015) Potential of Start Codon Targeted (SCoT) markers to estimate genetic diversity and relationships among Chinese *Elymus sibiricus* accessions. Molecules 20(4):5987–6001

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Turkish Journal of Physiotherapy and Rehabilitation; 32(3)
ISSN 2651-4451 | e-ISSN 2651-446X

BLOCKCHAIN TECHNOLOGY IN PREVENTING THE THEFT OF HUMAN ORGAN AND RARE BLOOD GROUP

Yogesh Sharma[1], B. Balamurugan[2]
[1]*Research Scholar, Galgotias University, India*
[2]*Professor, Galgotias University, India*
Yogeshsharma027@gmail.com, *bbalamurugan@galgotiasuniversity.edu.in*

ABSTRACT

Blockchain technology is an emerging technology that came into existence with the purpose of cryptocurrency but as the technology is entering into different sectors the technology is giving multiple benefits to the users of the blockchain. Blockchain technology is temper-proof and one of the most secure technology to date. These features of the technology have made the technology be used in many different sectors where the security and privacy of the information is a major concern. The Healthcare sector is one such sector where blockchain technology has played a very important role in securing the information of any patient and providing privacy to the health-related problem of any patient. One such case is to protect against the theft of human organs and rare blood groups. In most countries especially in India, there are numerous cases of organs and rare blood groups being stolen from the organ donation camps and the blood Banks respectively. The reason for this varies from inappropriate security measures taken to transaction failures like dual transactions. With this project such transactional and security concerns could be controlled thus preventing the theft to a great extent.

Keywords: Blockchain, Cryptocurrency, Security, Privacy, Healthcare, Human Organ.

I. .INTRODUCTION

The blockchain is an undeniably ingenious invention [1]. It is the brainchild of a person or a group of people known by the pseudonym Satoshi Nakamoto but since then it has evolved into something greater. Blockchain technology came into existence for the cryptocurrency called Bitcoin. However, as time progressed the tech community found that technology could be useful in many other sectors. At present, the total worth of the popular cryptocurrency Bitcoin is nearby around $150 billion [2] [3]. Similar to the internet a person doesn't need to know how the blockchain works to use it. However, only the elementary knowledge of this new technology can be useful to tell why this technology is considered an innovative technology. The information stored on a blockchain network exists as a shared and constantly submissive database. The database of the blockchain is not stored in a centralized location, which means that the records present in the database are public and could be easily verifiable kept by millions of computers at the same time. Its data is accessible to anyone on the internet. Blockchain technology is purely built upon the idea of distributed ledger technology in which each node on the network keeps a copy of the ledger [4]. Furthermore, the blockchain network follows a consensus, which means the network check for itself automatically after every 10 minutes [5]. Multiple nodes connected in a network create a blockchain. Every node in the blockchain has a complete copy of the ledger and joins the network willingly. Blockchain is a complex technology that is used to implement forgery-resistant chains [6] with a decentralized e-ledger, or in technical terms, it is a data structure for implementing chains with a unique ability to resist tampering [7]. As in the case of conventional database systems, transactions of blockchain can't be updated/removed, or inserted on the fly [8].

The ideology for implementing such a decentralized chained data structure was to make a transaction system with its currency, having the capability to retain information related to its previous transactions. Blockchain is now a catchphrase in business organizations and academia as well [9][10]. Immutability and temper resistance present the intrinsic value to cryptographic transactions by their ability to maintain true values globally and without any

jurisdiction [11]. Esurance of a blockchain system is administered by the security standards, which assure users of their sensitive data.

Blockchain is all about having secure and decentralized transactions. In most countries especially in India, there are numerous cases of organs and rare blood groups being stolen from the organ donation camps and the blood Banks respectively. The reason for this varies from inappropriate security Measures taken to transaction failures like dual transactions. With this project such transactional and security concerns could be controlled thus preventing the theft to a great extent.

II. .BLOCK

The blockchain, fittingly, is a chain of blocks and these blocks are the core of blockchain.

Each block consists of the following specific meta information

Block Data: Block Data comprises the data to be stored in encrypted format in the block along with the information of certified and authorized transactions tallied.

Block Header itself consists of the following

Block Number: Block Number can be easily understood as the index of the block in the chain. It is very crucial not to confuse the block number as the unique identifier of the block. The block number is directly proportional to the size of the blockchain [12].

Previous Hash: Previous Hash is the hash value of the previous block, mentioned in the current block to establish a link. The previous hash value prevents tampering. Previous Hash for genesis block is 0: The genesis block is the first i.e. base block of the blockchain.

Current Hash: Current Hash is the generated output of the cryptographic hashing function which encrypts the input values to a specific length. Generally, the input provided to the hashing function consists of the block's index, data, previous hash, timestamp, and the nonce value.

Timestamp: Timestamp refers to the time of block creation.

Size of the block(optional): The size of the block is its memory related to previous transactions retaining capacity.

Nonce value: The nonce value is generally referred to as the measure of difficulty, it an acronym for 'number only used once. The nonce is the value blockchain miners are mining for. The nonce value is hashed in the block which when rehashed, puts constraints on the mining complexity.

Authenticity and validation of block guarantee the expected digital signature i.e. previous hash. A non-validated and non-authenticated block would not be appended to the chain.

LEDGER of blockchain is a digital journal, distributed among all its users. It maintains a record of all transactions in that particular blockchain network.

A recorded blockchain transaction can't be altered. The ledger record of the transaction consists of the public keys of users involved in the transaction. In the transaction, a user shares its sensitive information with a specific user. Blockchain technology promotes the conception and philosophy of ledger's distributed ownership [13] [14]. One can simply interpret blockchain technology as nothing but a robust and highly secure implementation of distributed e-ledger technology.

III. .TYPES OF BLOCKCHAIN

Blockchains are popularly classified or tagged as Public, Private and Hybrid blockchains. This segregation of various blockchains is based on their permission standard.

3.1 Public or Permissionless

Blockchains are decentralized, distributed, and open to the public i.e. anyone can publish a block shown in fig. 1. These Permissionless blockchain networks are accessible to all irrespective of their terrestrial location [15], all needed is, internet and required software to run the blockchain network. Besides, the authority of writing data onto the blockchain is available to all connected users, whose consent is a must before storing any data in the database (ledger). A popular model of a public blockchain is Bitcoin [16]. This e-currency empowers users to use a platform for conducting transactions directly between them without involving any mediator like banks etc. There are a few advantages and disadvantages of Permissionless Blockchain.

Advantages:

Globally access.

Truly decentralized and translucent.

Disadvantage:

The rules of blockchain written for security persistence are public, and hence malicious attacks are relatively more in these open blockchain networks, which further increases the necessity of high computational power to ensure security and authenticity.

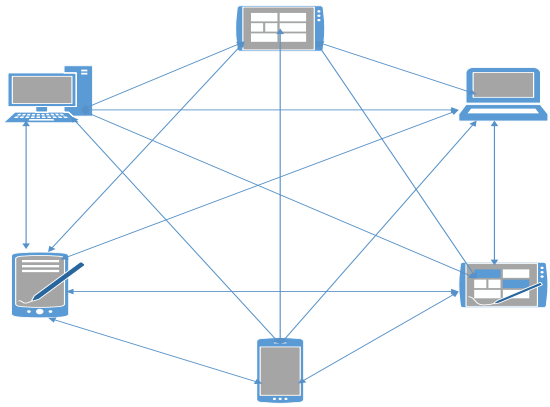


Figure 1. A Public Blockchain

3.2 Private or Permissioned

Blockchains are distributed chains having centralized or decentralized administration which sanctions the addition of the block into the chain and access of ledger. Transaction tracing is done through ledger records. The mining model used in private chains doesn't require the necessity of high computational power deployed to ensure authenticity and security as the number of malicious attacks is extremely less, besides, all the users' identities are directly or indirectly known to the members of this exclusive network. Consortium blockchain can be interpreted as a modified version of private blockchains [4]. In consortium chains, instead of a specific organization acting as the authorized body to control the entire processing and functionality of the chain, certain specific nodes i.e. users are given this authority which could be renewed periodically. Consortium blockchains are partially centralized, regardless of which they are popularly used and are more efficient than public chains. For instance, Permissioned blockchain systems are used by organizations that require a relatively more rigid and robust mechanism to secure their (exclusively) shared information. However, if an individual regulates the publishing of blocks, the members demand to have trust in the individual; permissioned blockchain networks can also be implemented in organizations that prefer, wishes or are required to work together but do not entirely trust each other with sensitive data.

There are a few advantages and disadvantages of Permissionless Blockchain.

Advantages:

Less computational power.

Lesser malicious attacks.

Faster than public blockchains.

Disadvantage:

The Private or Permissioned type of blockchain is partially transparent since only a few users of the group or a specific user have unrestricted access to the chain, which could lead possible chance of information theft as shown in fig.2.

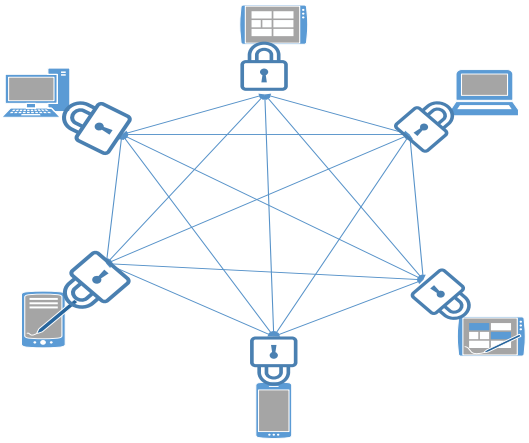


Figure 2. A Private Blockchain

3.3. HYBRID

Blockchains combine the privacy privilege of a private blockchain with the security and transparency privileges of a public blockchain. The blend of both public and private blockchain, the flexibility to health-based organizations, for instance, the freedom to maintain both data segmented as public and private information respectively. Hybrid chains join themselves with other private and public blockchains which expands the usage of the hybrid blockchain [17]. A well-recognized hybrid blockchain is also known as 'Dragon Chain'.

IV. IOT DEVICE

IoT is a network of "things" that are embedded with sensors, software, and other technologies that are connected via any network (generally the internet) to exchange data and control from any remote location. The IoT devices have to be created with a unique identifier, these devices have the ability by which any data can be transferred over a network that does not require any human or any node interaction. The data is pushed onto the internet/Bluetooth via the ESP32 chip integrated with the sensors, actuators, and/or embedded system. Esp32 uses a TCP Server as a part of the software that will run. The job of Esp32 is to look for the incoming connections at a port on the device's IP address. It is when a remotely situated computer, or client, reaches out to the TCP Server it generates a channel for the ESP32 and the client at the remote site for the communication. An Arduino program for the ESP32 uses a WiFiServer object to generate a server. Whenever an original connection is created, the server delivers a WiFiClient object as the communication channel. ESP32 further uses Wi-Fi Protected Access (WPA) as a security standard for the devices equipped with wireless internet.

In the present scenario, a NodeMCU (uses ESP32 chip) is connected to a keypad and a servo motor, thus making it a node, connected to the internet using a secure network. As the user enters their phone number, it is uploaded to the cloud, through which it is loaded to the server generating a unique password for each user which is sent to their phone via a message. Once the number is verified, a signal is sent from the popularly used in the cases where there is a requirement of generating irreversible hash statements. It is very important to understand the difference between hashing and encryption, a hashed value cannot be used to regenerate the original content from which it was derived, whereas an encrypted value can be decrypted back to its original value if the required values like salt and secretKey are known.

example:

SHA256("")

0x e3b0c44298fc1c149afbf4c8996fb92427ae41e4649b934ca495991b7852b855

Original Text: Hello, how are you, I have high cholesterol value.

SHA256Value: 1fa6e4f59486f72e3e504cfca9981416297213541967c2c4682791e161106be0

Modified Text: Hello, how are you, I have a low cholesterol value.

SHA256Value: 26bccd725246b89cbdd74cca64014ed399ccf0c24011068ab89a9eabf4dcef5b

Here we have seen modifying small text in the document could modify the complete hash value of the document. This shows how important is the hash algorithm in data security and privacy.

V. RSA

RSA is an abbreviation for Rivest–Shamir–Adleman. It is an asymmetric cryptographic algorithm that is widely used for the encryption and decryption of messages [18]. It returns a public key and a private key. RSA is based on the fact that it is relatively difficult to factorize a large integral value [19] [20]. The public key comprises of two numbers, one of the numbers is a multiplication of two relatively large prime numbers[20], similarly, the private key is also derived from the same two prime numbers [21]. Thus, to compromise the private key, one has to factorize a large prime number. Therefore, encryption durability depends on the key size i.e. length of the key and if one doubles or triples the size of the key, the strength of encryption will increase exponentially making the strength and size directly proportional to one another.

VI. KEYS

There are two types of keys used in the blockchain: Public key as the name suggests could be shared publicly or in general, the ledger of blockchain stores its entries against the public key of the user, whereas the private key is supposed to be maintained as a secret. The public key is used for encryption of the message or the data to be shared, and the private key is used for the decryption of this data. Public Key and Private Key both keys are generated using algorithms like RSA. A private key is generally 256 bits whereas a public key is 2048 bits. Because of having a smaller size, private keys are faster than public keys.

VII. .SYSTEM ARCHITECTURE

7.1 Architecture Block Diagram 1 Authentication

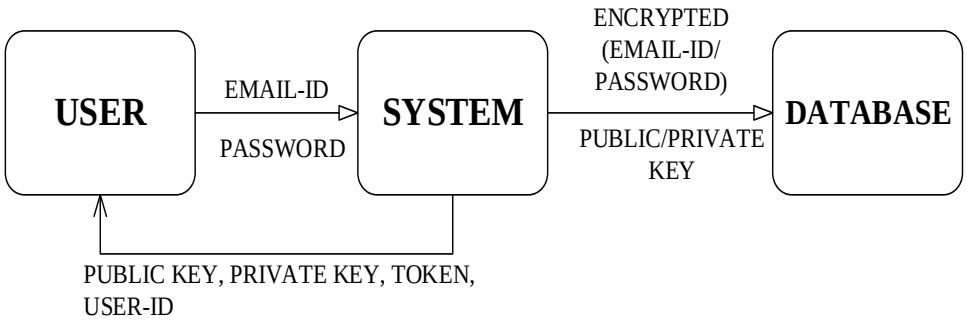


Fig.3 Authentication Process

The user will authenticate using an email and a password; these both will be encrypted and stored in the database along with generated set of keys i.e. the public key and the private key which will be given to the user along with the token.

7.2 Architecture Block Diagram 2: Creating Block

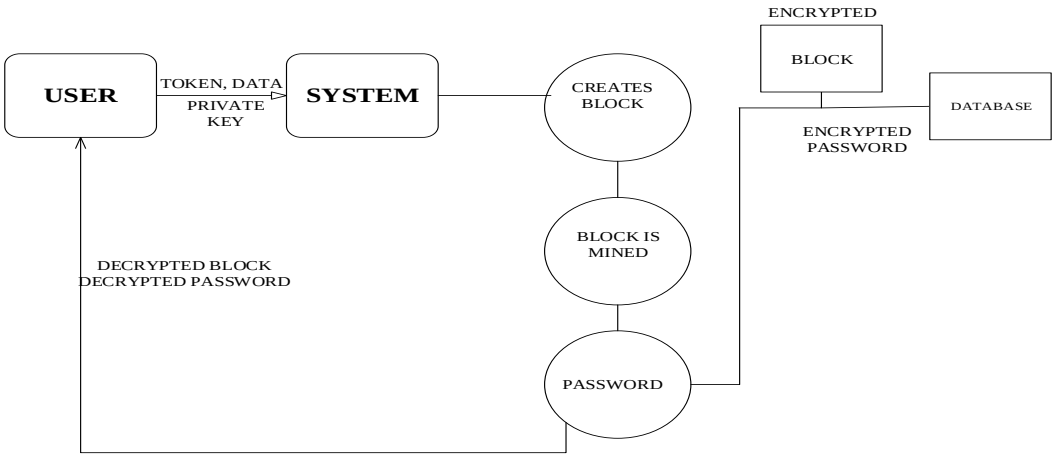


Fig.4 Creation of Block for Blockchain

User will provide data for transaction i.e. requirements of the Organ(s)/unit(s) of blood group along with a private key. The application will automatically send a private key and authentication token to the backend server which will verify the private key to the token's corresponding user's private key. If they match, the system will create a block that will have ID, index, previous hash, current hash, the data, and nonce.

(note: Previous hash is zero for the Genesis block and hash generated by mining, act like the previous hash for the subsequent blocks)

On successful creation of a block i.e. after mining, a block will be attached to the chain.

Subsequently, the password will generate and store in the block's DB in encrypted format along with encryptions of data and the creator's email. the password will be shown to the user along with other related information of the block in decrypted format at the time of creation.

8.3 Architecture Block Diagram 3: Public Ledger

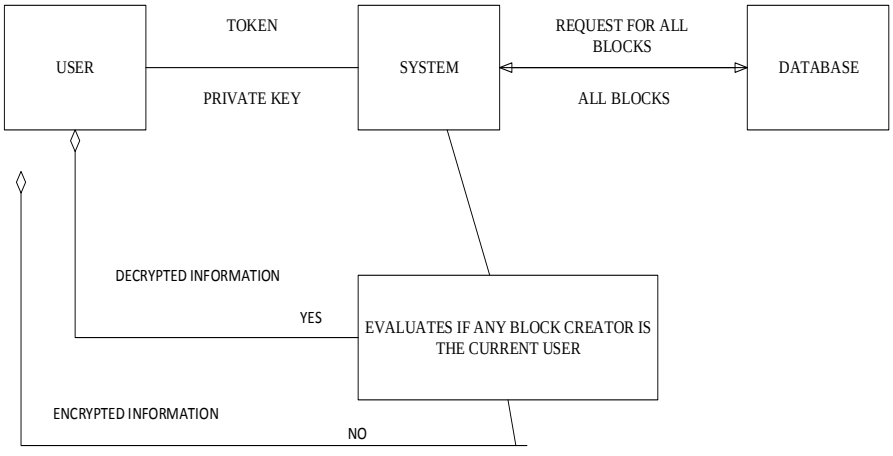


Fig. 4 Request to Public Ledger

The user will provide a private key and the application will send it along with an authentication token to the backend system. After successful verification of keys, information of the ledger will be displayed to the user. Meanwhile, an asynchronous process will run to cross-check for all the blocks if their creator and accessing user are the same then decrypted password corresponding to those particular blocks will be shown else all the information will remain encrypted.

8.4 Architecture Block Diagram 4: Request Access of Safe

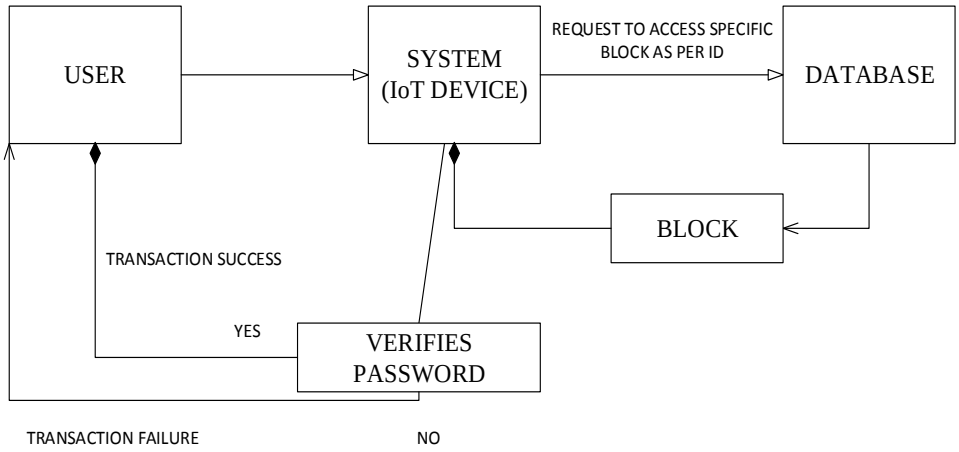


Fig. 5 Request to access safe

The user will input his authentication credentials and block ID into IOT Device which will be connected to the backend server, sending a post request for entered credentials to the respective API endpoint.

The system would verify credentials and will provide access to blockchain only to the authorized user, followed by the block ID entered being verified against all blocks created by the current user, if a match is found then the user is requested to input the password corresponding to that particular block, if this password and the one generated at the time of block creation (stored in encrypted format matches), then the user will be given full access to read decrypted information of transaction/access the safe, else transaction would be declared unsuccessful.

Note:

1.The credentials of the block are stored in a database in an encrypted format.

2. Only three failed transactions per block ID are valid else the block would be permanently marked for unauthorized access thereby preventing a security breach.

VIII. .IMPLEMENTATION

Block is a class/constructor function that takes in 3 parameters

Index: The index of the block

Data: The data to be stored inside the block

PrevHash: The hash of the previous block in the chain

The constructor function handles variable value assigning, and also creating the timestamp value, and calculating the hash for the block.

The timestamp is assigned using the Date. Now, the function of JavaScript, and for ease of working, nonce value is hardcoded to 1.

clash function uses:

- index
- timestamp
- data
- prevHash

- nonce

to create the hash value for the current block using the SHA256 algorithm.

For ease of workflow, custom mining action is attached to the block which carries out the mining action as shown in the script.

```
class Block {

  constructor({index, data, prevHash}) {
     this.index = index;
     this.data = data;
     this.prevHash = prevHash;
     this.timestamp = Date.now();
     this.nounce = 1;
     this.hash = this.calcHash();  }

  calcHash(): string {
     return sha256(
    this.index + this.timestamp + this.data + this.prevHash + this.nounce
     )
  }

  mineBlock(difficulty: number): string {
     while (this.hash.substring(0, difficulty) !==       Array(difficulty + 1).join("0")) {
        this.nounce++;

        this.hash = this.calcHash()

     }  return this.hash

  }

}
```

The created block function takes in the following parameters

- index
- data
- prevHash

On passing the following information to the Block class, a custom block is returned, and it is stored in a variable: toAddBlock.

if prevHash i.e. previous hash, is 0 then the block is returned as the genesis block, else the mining action takes place to ensure the validity of the information passed in the block, if the information is valid, then the block is returned else **false** is return, which means that the block is invalid.

```
type TCreateBlock = {
  index: string;
  data: string;
  prewash: string
```

```
}
const createBlock = (blockInfo: TCreateBlock) => {
  const toAddBlock = new Block(blockInfo)
  if (blockInfo.prevHash === "0") {
     // genesis block
     return toAddBlock
  } else {
          // mineBlockTakes in difficulty value

     const returnedHash = toAddBlock.mineBlock(2);

     return returnedHash === toAddBlock.hash ? toAddBlock : false

  }

};
```

IX. .CONCLUSION

There is diverse work being done around the globe for expanding the horizon of blockchain. Though it has been used more than just as a financial tool, also in healthcare and security besides numerous others which haven't yet come into the picture. The current idea of using blockchain as another security measure for generating passwords and ensuring secured transactions of extreme essential commodities specifically like human organs and blood groups. By using such a mechanism, the theft could be brought down to a relatively lesser number.

Yes, various factors are limiting the expanding horizon of the blockchain specifically its time complexity and mining but there is no denial of the fact that one day soon, industries and organizations around the globe would be using the power of blockchain in their transaction systems which may belong to security, financial or document.

REFERENCES

1 Politou, E., Casino, F., Alepis, E., & Patsakis, C. (2019). Blockchain mutability: Challenges and proposed solutions. IEEE Transactions on Emerging Topics in Computing.
2 Conti, M., Kumar, E. S., Lal, C., & Ruj, S. (2018). A survey on security and privacy issues of bitcoin. IEEE Communications Surveys & Tutorials, 20(4), 3416-3452.
3 Mohamed Rahouti, K. X. (2018). Bitcoin Concepts, Threats, and Machine-Learning Security Solutions. Mohamed RahoutiKaiqi XiongKaiqi XiongNasir Ghani, PP (99):1-1.
4 Yogesh Sharma, B.Balamurugan (2020). A survey on privacy-preserving methods of Electronic medical record using blockchain". journal of mechanics of continua and mathematical sciences, pp 32-47.
5 Wang, W., Hoang, D. T., Hu, P., Xiong, Z., Niyato, D., Wang, P., & Yonggang Wen. (2019). A Survey on Consensus Mechanisms and Mining Strategy Management in Blockchain Networks. IEEE Access, 22328 - 22370.
6 Mosakheil, J. H. (2018). Security Threats Classification in Blockchains. St. Cloud State University
7 Massimo Bartoletti, S. L. (2017). A general framework for blockchain analytics. Proceedings of the 1st Workshop on Scalable and Resilient Infrastructures for Distributed Ledgers (pp. 1–6). ACM.
8 Ganne, E. (2018). Can Blockchain revolutionize international trade?". Switzerland: World Trade Organization.
9 Pilkington, M. (2016). Blockchain technology: principles and applications. In Research handbook on digital transformations. Edward Elgar Publishing.
10 Vera Maltseva, A. M. (2019). Blockchain and the Future of Global Trade (Review of the WTO report “Can Blockchain revolutionize international trade”). INTERNATIONAL ORGANISATIONS RESEARCH Journal, 191-198.
11 Zhou, E., Sun, H., Pi, B., Sun, J., Yamashita, K., & Nomura, Y. (2019, October). Ledger data refiner: a powerful ledger data query platform for hyperledger fabric. In 2019 Sixth International Conference on Internet of Things: Systems, Management and Security (IOTSMS) (pp. 433-440). IEEE.
12 Atzori, M. (2015). Blockchain technology and decentralized governance: Is the state still necessary? Available at SSRN 2709713.
13 Bryanov, K. (2018). Distributed Ledger of Ideas: Toward A Blockchain-Powered Marketplace ‘RAP on Blockchain’ at North Dakota State University's Center for the Study of Public Choice and Private Enterprise.
14 Ge, C., Sun, S., & Szalachowski, P. (2019, May). Permissionless Blockchains and Secure Logging. In 2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC) (pp. 56-60). IEEE.
15 Rui Zhang, R. X. (2019). Security and Privacy on Blockchain. CSUR 2019. ACM.
16 T.K.Sharma. (2018). Advantages and Disadvantages of Permissionless Blockchain, Retrieved from Https://www.Blockchain-Council.Org/.
17 Gokhan Sagirlar, B. C. (2018). Hybrid-IoT: Hybrid Blockchain Architecture for Internet of Things - PoW Sub-blockchains. IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), (pp. pp. 1007-1016). Halifax, NS, Canada, Cybermatics_2018.2018.00189.
18 Rohit Minni, K. S. (2013). An algorithm to enhance security in RSA. Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), 1-4.
19 Siahaan, A. P. (2016). Factorization Hack of RSA Secret Numbers. International Journal of Engineering Trends and Technology. 15-18.

20 Muhammad Ariful Islam, M. A. (2018). A Modified and Secured RSA Public Key Cryptosystem Based on “n” Prime Numbers. Journal of Computer and Communications 06(03): 78-90.

21 Mustafa M. Abd Zaid, D. S. (2018). Lightweight Rsa Algorithm Using Three Prime Numbers. International Journal of Engineering and Technology (UAE), 293-295.

12.Authors Profile



Prof. B. Balamurugan, Completed his Ph.D. at VIT University, Vellore, and currently working as a Professor at Galgotias University, Greater Noida, Uttar Pradesh. He has 15 years of teaching experience in the field of computer science. His area of interest lies in the field of the Internet of Things, Big data, Networking. He has published more than 100 international journal papers and contributed book chapters.

Email: bbalamurugan@galgotiasuniversity.edu.in



Mr. Yogesh Sharma is a research scholar in the School of Computer Science and Engineering, Galgotias University, Greater Noida, Uttar Pradesh, India. He has completed ME in Computer Technology and Application, from Delhi Technological University (Formerly Delhi College of Engineering), Delhi, India. His interest lies in the field of Big Data, Data mining, Network security, and Blockchain. He has published 10 papers in various international conferences and journals. He has two patents in his name. He has also written book chapters in reputed publications.

Email: Yogeshsharma027@gmail.com

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/351375593

Capacitive Investigation of a newly introduced TF-FinFET

Research · May 2020

CITATIONS
0

READS
41

2 authors:



Mridul Prakash Kashyap
Delhi Technological University
6 PUBLICATIONS **5** CITATIONS
SEE PROFILE



Rishu Chaujar
Delhi Technological University
218 PUBLICATIONS **1,321** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Modelling and Simulation of GAA FinFET View project

Introducing Hybrid Transistors for Robust Wireless applications View project

All content following this page was uploaded by Mridul Prakash Kashyap on 29 May 2021.

The user has requested enhancement of the downloaded file.

Capacitive Analysis of a Novel 3 nm Truncated Fin FinFET

Mridul Prakash Kashyap[1] and Rishu Chaujar[2]

Delhi Technological University, New Delhi. 110042, Delhi, INDIA
[1]mridul.prakash2012@gmail.com,[2]chaujar.rishu@dtu.ac.in

Abstract. In concerned work, a novel *Truncated Fin* (TF) FinFET is studied for its 3 different oxide thicknesses which are solely compatible to its gate length. All in all, the analysis of the extra-ordinarily switching performance has been conducted for a JL-bulk FinFET at the gate length of 3 nm, with the help of Atlas Silvaco designing software. Additionally, the capacitance dependent figure of merits such as Gain Bandwidth Product (GBP), Transconductance, and Energy Delay Product (EDP) are also analysed and found that, trade-offs between bandwidth and power are minimum at Tox=1 nm in TF-FinFET which ultimately dedicated to a noticeable reduction in parasitic capacitance (Cgg = Cgs + Cgd). This is because EDP and GBP also improve consequently and hence will reflect its effectiveness from RF amplifiers and receivers' point of view. As TFP and EDP of TF-FinFET at Tox=1 nm are only 38.4% and 43.5% of TFP and EDP respectively at Tox=0.4 nm. Hence device is most efficient at Tox=1 nm. Extracted results from simulation showed the optimized characteristics of TF-FinFET for its different oxide thicknesses, from the perspective of system on chip (SOC) applications.

Index terms: Truncated fin FinFET, Parasitic capacitance, TCAD, Capacitive FOMs

1 Introduction

As suggested by Moore's Law, the size of transistors is decreasing intensively due to the sophisticated designs come into picture. From last two decades approximately, the transistors have been reasonably scaled down in terms of dimensionality to accomplish high circuit density and improved performance in a specified chip constraint. As the channel region is in contact with the substrate (GaAs) and having an electrode connected to its truncated part at the same time. It bears all the advantages of bulk FinFET with improved performance, as only one-third of conventional substrate width is dedicated to its bulk structure and two-third of fin width is dedicated to truncated structure mechanism [1-3].

But, the capacitive role of the device is always being a matter of concern when we implement a transistor in VLSI and ULSI methodology. Capacitors are always considered as parasitic as they affect the switching speed of the device. They come into the picture when input of the device goes from high-low or low-high [4]. Amount of capacitance is inversely proportional to the switching speed. More the capacitor in our

device, more will be the energy storage in it in the form of charge and hence more transit time will be require to completely charge or discharge a capacitor. If this property of a transistor is seen collectively, then there would be large delays in our digital systems [5]. The main novelty in our device is providing combined characteristics of GAA and 3-gated FinFETs simultaneously, due to the truncatedness and material (i.e., GaAs) of substrate used. But here, we discussed its Capacitive-Voltage (C-V) and Capacitive-Frequency (C-F) characteristics as a function of its different width of gate oxides solely [6-8].

2 Description of Device Structure

Fig.1. (a) and (c) shows the truncated fin structure and conventional structure respectively, and (b) shows the more described capacitance spread in TF-FinFET from its 2D view.

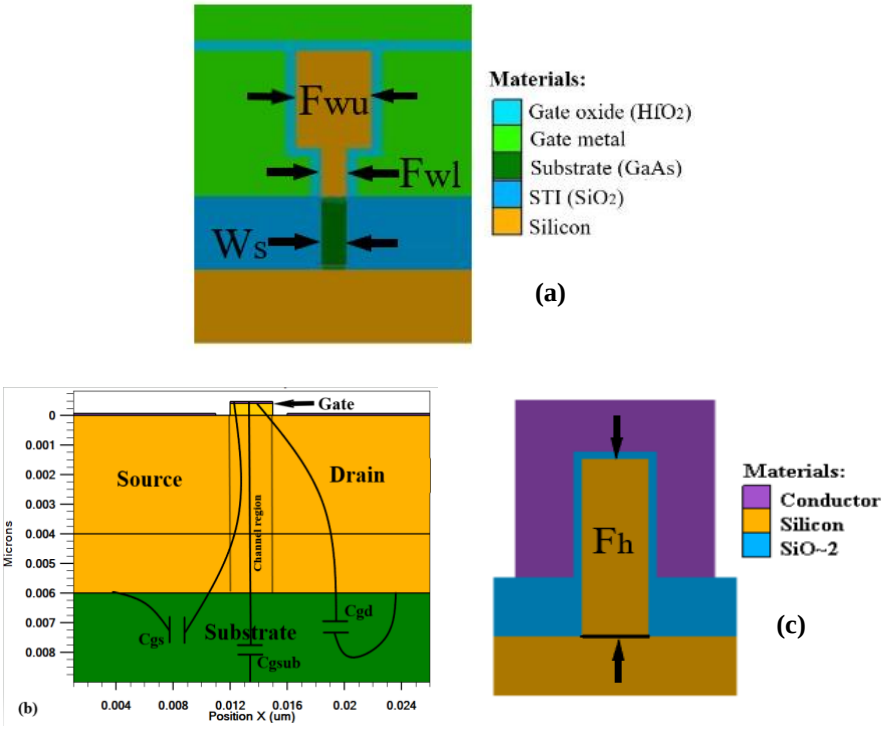


Fig. 1. a) 2-Dimensional cross-sectional representation of JL bulk TF-FinFET. b) Representation of parasitic capacitances in TF-FinFET after bisecting it. c) 2-Diensional cross-sectional representation of JL conventional FinFET.

The oxide thickness is varied from 0.4 nm to 1 nm in step of 0.3 nm. The gate oxide material for TF-FinFET has been chosen as HfO_2 as compared to SiO_2 in given

conventional FinFET, which as a result optimizes the gate controllability and hence the amount of driving current increases during conduction. The overall length and height of the device has been restrained to 11 nm and 9 nm respectively. As far as doping profile is concerned, uniform doping has been taken into account for all the regions such as $1x10^{20}$ (cm^{-3}) for source, $1x10^{18}$ (cm^{-3}) for drain, $1x10^{19}$ (cm^{-3}) for channel and $1x10^{15}$ (cm^{-3}) for substrate. The doping profile was set to provide the fabrication ease for manufacturer point of view. Moreover, the truncated part is somehow used to bring the GAA properties in the conventional 3-gated FinFET [10] and remaining parameters are shown with the help of table.1.

Table 1. Structural dimensions for simulated device with different Tox

Dimensional Quantities	**Values of Dimensions (n-type)**
Gate length (L_G)	3 nm
Fin height (F_h)	6 nm
Lower fin width (F_{wl})	1 nm
Upper fin width (F_{wu})	3 nm
Gate oxide thickness (Tox)	0.4,0.7,1.0 nm
Gate metal workfunction	4.55 eV
Contact resistivity	417 nΩ.m
Substrate width (W_S)	1 nm

3 Device Simulation Methodology

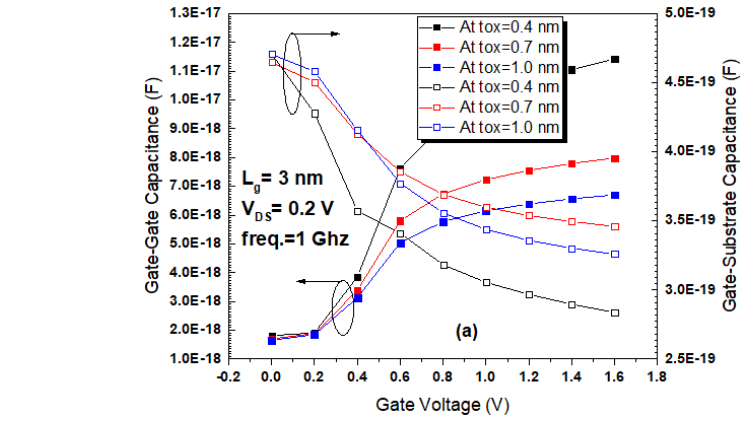
All 3-D Simulations are simulated on Atlas Silvaco TCAD Tools. As Table I. Shows structural parameters of the proposed TF-FinFET. The fin has been assumed as a combination of 2 sub parts (i.e., upper fin and lower truncated fin). Lower truncated fin is designed so as to remain its thickness one-third of the upper fin thickness. As a result, this adjustment actually enhanced the gate controllability over the channel region. Consequently, as shown from the same table, the substrate width is also made thinner by 2 nm. The material of substrate is chosen as GaAs [11-12].

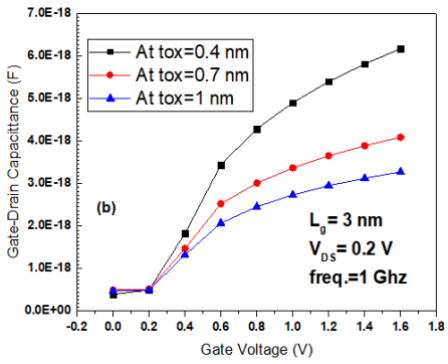
If seen from fabrication perspective, the appropriate physical models for quantum correction has been used for predicting the behavior of our TF-FinFET, therefore its actual implementation onto the circuit board. These are namely concentrated dependent mobility to examine the mobility of charge carriers on different doping profiles, SRH recombination model to predict the discrepancy occurred in electrical transfer characteristics due to interaction of majority charge with immobilized ionized atoms in

channel region, Fermi-Dirac statistics model to analyze the distribution of identical particles that obeys the Pauli exclusive principle and also dedicated to distort the input characteristics of our device, Hydrodynamic energy transport model to treat the mobility and impact ionization coefficients as a function of temperature and Band to Band tunneling model to examine the charge carrier movements in between the bands that shouldn't happen, if seen from perspective of classical theory.

4 Results of Simulation

In concerned article, all simulations are carried out for n-type device. The truncation in TF-FinFET actually restricted the mobility of electron charges up to upper fin region to avoid leakage current at ON state through the substrate (GaAs) as much as possible and that's the main reason, why TF-FinFET has somehow a little bit higher capacitance as compared to its conventional structure [13-15]. But this drawback can easily be compensated by its performance enhancement from conventional counterpart, if seen analog point of view.





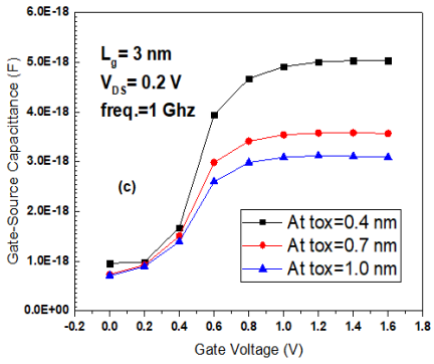


Fig. 2. a) Dependency of gate-gate and gate-substrate capacitances on V_{GS}. b) Dependency of gate-drain capacitance on gate to source voltage. c) Dependency of gate-source capacitance on V_{GS}.

As shown with fig.2(a), the variability of gate-gate capacitance (C_{gg}) and gate-substrate capacitance (Cgsub) on gate voltage (V_G) and applied input frequency of 1 GHz is studied. The variance of C_{gg} analyzed with the concerned plot representation, is actually a consequence of the gathering of electrons in the proximity of gate region and hence the parasitic (Cgg = Cgs + Cgd) capacitance rises. If seen from the perspective of oxide thickness, C_{gg} stood maximum for Tox=0.4 nm. This is due to the basics reason of inverse dependency of capacitance on the plate's separation. In case of C_{gsub}, the formation of shielding layer (electron charge carriers) takes place in channel region between gate and substrate [16]. This phenomenon is wholly responsible for the variation of C_{gsub} w.r.t gate voltage as shown. C_{gsub} is minimum in case of FinFET with Tox=0.4 nm, as accumulation of charge carriers to the substrate is most restricted due to least leakage of drain current and hence having best gate controllability [17].

In context of Fig.2(b), we can depict that gate voltage actually increases the drain current through channel. These current passes more frequently and with more intensity, due to which C_{gd} increases with increment in gate voltage. The minimum capacitance can be analysed from the respective fig. for Tox=1 nm. As in this case, the gate controllability is least and less transfer of electrons from source to drain takes place. Similarly, in case of Fig.2(c), the magnitude of gate-source capacitance is minimum for Tox=1 nm. From V_{GS}=0.7 V onwards, the saturation of drain current (I_D) takes place [18]. This phenomenon can easily be depicted from the charge accumulation due to which the C_{gs} and C_{gd} also tends to get saturate. The charge accumulation is somehow greater near drain region due to the occurring of pinch-off point and formation of reverse biased junction Between drain and substrate and hence C_{gd} is greater than C_{gs} by a very small amount.

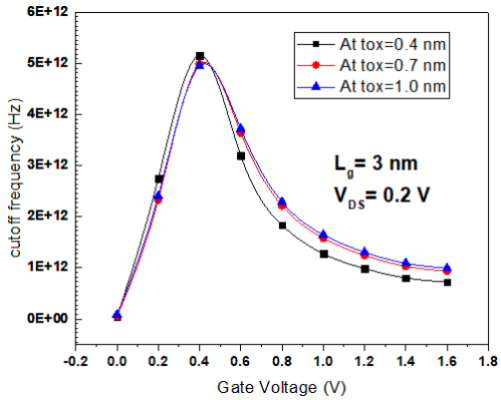


Fig. 3. Dependence of cutoff frequency on Gate Voltage at drain voltage of 0.2 V.

As f_T is described as the input frequency from which current gain starts to owe a unity value. One can easily take idea from the respective figure that the current is being started to saturate from V_{GS}=0.4 V onwards. That is how we obtained a peak in fig.3 at V_{GS}=0.4

V. Cutoff frequency is maximum for FinFET with Tox=0.4 nm because of better gate controllability in this case.

Further the variation of C_{gg}, C_{gsub}, C_{gd} and C_{gs} is observed by varying the frequency of input signal at gate terminal. It is clearly reflected from the respective graphs that all the capacitances remained quit stable from MHz to THz range of frequency. Above this range they started varying more vulnerably, which ultimately conclude the stability of this device upto THz range of input signal frequency. C_{gg}, C_{gd} and C_{gs} are observed minimum in case of with Tox=1 nm, occurred due to the minimum controllability over gate and less transfer of electrons from source to drain takes place in this case. As far as C_{gsub} is concerned, it increases as frequency of input signal increases [19]. The major reason responsible for this phenomenon is decrement in the shielding between gate and substrate. It is min. for Tox=0.4 nm because of maximum shielding in this case.

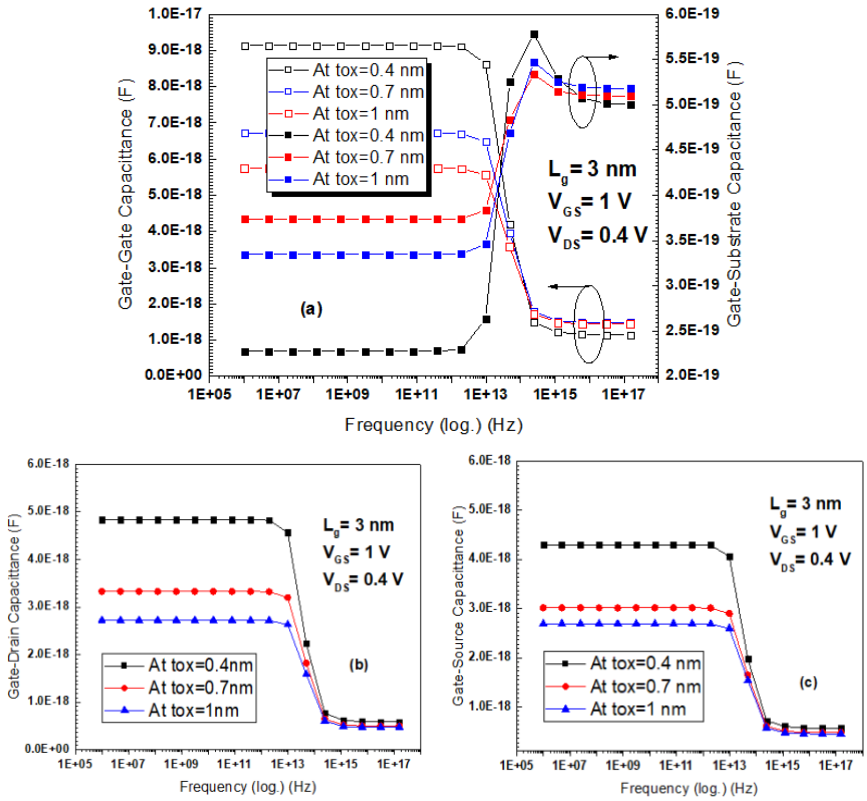


Fig. 4. a) Dependency of gate-gate and gate-substrate capacitances on frequency. b) Dependency of gate-drain capacitance on frequency. c) Dependency of gate-source capacitance on frequency. At V_{DS}=0.4 V and V_{GS}=1 V.

Some of the few important FOMs are shown below which describes the high frequency performance of the device.

$$TFP = (g_m/I_D)(g_m/2\pi C_{gg}) \quad (1)$$

$$EDP = (C_{gg}V_D^2)(C_{gg}V_D/I_{on}) \quad (2)$$

$$GBP = (g_m)/(2\pi \times 10 \times C_{gd}) \quad (3)$$

TFP (Transconductance frequency product) is majorly the product of device efficiency and cut-off frequency (f_T). It signifies the trade-off bandwidth and power which ultimate dedicated to high-speed designs [20]. It decreases significantly from V_{GS}=0.2 V onwards due to the increment in C_{gg} as shown in ***fig.5b***. It is observed that TFP is highest for Tox=0.4 V. This is due to the high value of transconductance in this case.

Energy delay product (EDP) is defined as the product of average energy (E_{ave}) and gate delay (τ). It refers the energy consumed in each switching cycle. In case of Tox=1 nm, it stood at lowest level as shown in ***fig.5a***. This is due to the lowest gate-gate capacitance in this particular case. Further the Gain Bandwidth Product (GBP) signifies the amplifier gain and corresponding bandwidth over a particular voltage range. It is maximum when Tox=1 nm, as C_{gd} is least in this case [22].

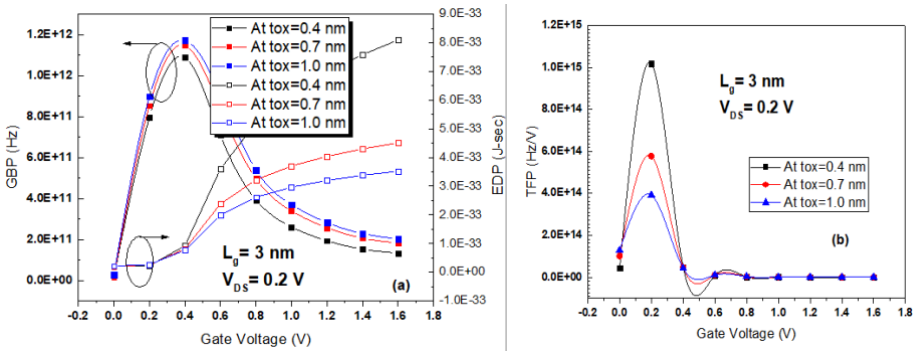


Fig. 5. a) Dependence of GBP and EDP on V_{GS}. b) Variation of TFP w.r.t V_{GS}.

4.1 Characteristics Difference between Conventional and TF-FinFET

TF-FinFET is compared with conventional one on the basis of its analog performance. All the structural and doping parameters kept same between the 2 devices excepting the

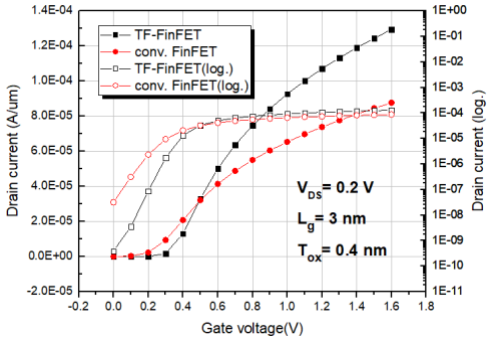


Fig. 6. I_D-V_G of the simulated devices at V_D=0.2 V

truncation in fin bottom, material of substrate and gate oxide.

Fig.6 shows Input characteristics of both the devices at particular constant drain voltages (i.e., V_{DS}=0.2 V). In case of V_D=0.2 V, the saturation current (i.e., at V_{GS}=1.6 V) comes out to be $1.29x10^{-4}$ A/fin for TF-FinFET and $8.76x10^{-5}$ A/fin for the conventional FinFET. The major difference is observed in their I_{OFF} (at V_G=0 V). I_{OFF} for TF-FinFET is found to be $3.91x10^{-10}$ A/fin and $3.18x10^{-8}$ A/fin for conventional FinFET [13-14]. As channel is more depleted in case of our proposed FinFET structure as compared to the depletion of channel in conventional FinFET at off state (V_G=0 V). This is because of enhanced gate control over the channel part.

4.2 Calibration of TF-FinFET with a fabricated device

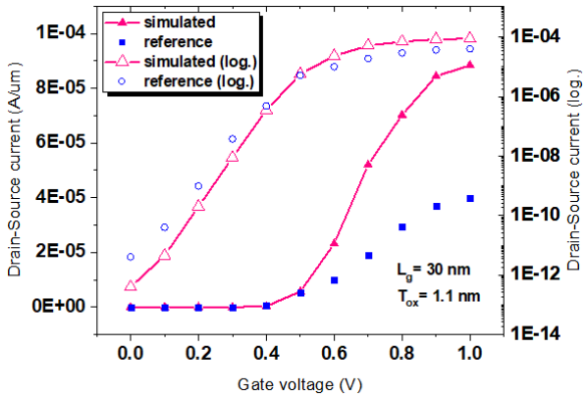


Fig. 7. Dependence of Drain current (I_D) on input gate voltage at V_{DS}=0.05 V.

Fig. 7 demonstrates the calibration of the models used during simulation of TF-FinFET with the input transfer characteristics of the fabricated devices at V_{DS}=50 mV. Here, from our convention, the curve of the fabricated device has been shown by "reference" and the simulated TF-FinFET by "simulated". These simulations are analysed at the scale of 30 nm for SiO_2 spacer by considering the paper [14], on Atlas Silvaco TCAD software. By referring the concerned device, all the simulated results stood quite comparable with those shown in paper [14]. Form here, one can easily depict the similarity in their respective input characteristics w.r.t ramping gate voltage (V_{GS}). Basically, plot showed the variation of drain current in Fabricated device, when simulated by using our models for the validation of our models mentioned in section III. Simulated device stood with similar to the results of fabricated device in terms of (I_D-V_G) characteristics.

5 Conclusion

This work is based on the analysis of Capacitance-Voltage and Capacitance-Frequency

characteristics of a novel 3 nm TF-FinFET. The impact of gate oxide widths on TF-FinFET is analyzed with respect to applied voltage and dependency of parasitic capacitances on corresponding frequencies and capacitive FOMs such as TFP, GBP and EDP for improved switching performance. All the (C-V) and FOMs analysis showed the performance enhancement of the device at Tox=1 nm. Hence it is most efficient at Tox=1 nm. TF-FinFET has somehow a little bit higher capacitance as compared to its conventional structure. But this drawback can easily be compensated by its performance enhancement from the conventional counterpart, if seen analog point of view. So, TF-FinFET stands very promising for the high frequency/switching applications at 3nm technology node. It has great a potential in high performance switching applications of SRAM and FPGA Designs**.**

Acknowledgement

This document is prepared in support of Microelectronics Research Laboratory, Department of Applied Physics, Delhi Technological University, New Delhi.

References

1. A. Kumar, M. M. Tripathi, and R. Chaujar, "Investigation of parasitic capacitances of In2O5Sn gate electrode recessed channel MOSFET for ULSI switching applications," *Microsyst. Technol.*, vol. 23, no. 12, pp. 5867–5874, 2017, DOI: 10.1007/s00542-017-3348-2.
2. A. Kumar, N. Gupta, and R. Chaujar, "Analysis of novel transparent gate recessed channel (TGRC) MOSFET for improved analog behaviour," *Microsyst. Technol.*, vol. 22, no. 11, pp. 2665–2671, 2016, DOI:10.1007/s00542-015-2554-z.
3. E. Yu, K. Heo, and S. Cho, "Characterization and Optimization of Inverted-T FinFET under Nanoscale Dimensions," *IEEE Trans. Electron Devices*, vol. 65, no. 8, pp. 3521–3527, 2018, DOI:10.1109/TED.2018.2846478.
4. J. Madan and R. Chaujar, "Numerical Simulation of N+ Source Pocket PIN-GAA-Tunnel FET: Impact of Interface Trap Charges and Temperature," *IEEE Trans. Electron Devices*, vol. 64, no. 4, pp. 1482–1488, 2017, DOI: 10.1109/TED.2017.2670603.
5. V. Thirunavukkarasu, Y. R. Jhan, Y. B. Liu, and Y. C. Wu, "Performance of Inversion, Accumulation, and Junctionless Mode n-Type and p-Type Bulk Silicon FinFETs with 3-nm Gate Length," *IEEE Electron Device Lett.*, vol. 36, no. 7, pp. 645–647, 2015,DOI:10.1109/LED.2015.2433303.
6. Bhattacharya, Monika & Jogi, Jyotika & Gupta, R. & Gupta, Mridula., "Scattering parameter based modeling and simulation of symmetric tied-gate InAlAs/InGaAs DG-HEMT for millimeter-wave applications", *Solid-state Electronics - SOLID STATE ELECTRON*, vol.63. pp.149-153, 2011, DOI:10.1016/j.sse.2011.05.025.
7. Singh, Ramendra & Aditya, Kritika & Parihar, Shivendra & Chauhan, Yogesh & Vega, R. & Hook, T.B. & Dixit, Abhisek, "Evaluation of 10nm Bulk FinFET RF Performance - Conventional vs. NC-FinFET", *IEEE Electron Device Letters*, vol.39, No.8, pp. 1246-1249, 2018, DOI:10.1109/LED.2018.2846026.

8. Sanchez Esqueda, Ivan, “Confinement Effects on Radiation Response of SOI FinFETs at the Scaling Limit”. *IEEE Electron Device Letters*. Vol.38, No.3, pp. 306-309, 2017, DOI:10.1109/LED.2017.2661239.
9. Zhang, Jiabi & Niu, Guofu & Cai, Will & Wang, Weike. “Intermodulation Linearity Characteristics of 14-nm RF FinFETs”, *IEEE Transactions on Electron Devices*. Vol.66, No.6 pp. 2520-2526, 2019, DOI:10.1109/TED.2019.2912516.
10. P. Zheng, D. Connelly, F. Ding, and T. J. K. Liu, “Simulation-Based Study of the Inserted-Oxide FinFET for Future Low-Power System-on-Chip Applications,” *IEEE Electron Device Lett.*, vol. 36, no. 8, pp. 742–744, 2015, DOI: 10.1109/LED.2015.2438856.
11. P. H. Su and Y. Li, “Analysis of In-Line Process Parameters of the Unity Gain Frequency of HKMG Bulk FinFET Devices,” *IEEE Electron Device Lett.*, vol. 39, no. 3, pp. 335–338, 2018, DOI: 10.1109/LED.2018.2791436.
12. ITRS 2.0 Publication, “International Technology Roadmap for Semiconductors, 2015,” Available online at http://www.itrs2.net/.
13. PH Drive, S. Clara, Atlas User’s Manual, Device Structure Editing and Simulation Software; 2017. Pp. 1-1715. Available: www.silvaco.com.
14. A. B. Sachid, M. C. Chen, and C. Hu, “Bulk FinFET With Low-κ Spacers for Continued Scaling,” *IEEE Trans. Electron Devices*, vol. 64, no. 4, pp. 1861–1864, 2017, DOI:10.1109/TED.2017.2664798.
15. Bolognesi, Colombo & Dvorak, M.M.W. & Yeo, Philip & Xu, X.G. & Watkins, S..” InP/GaAsSb/InP double HBTs: a new alternative for InP-based DHBTs”. *IEEE Transactions on Electron Devices*. Vol.48. pp. 2631 - 2639. DOI: 10.1109/16.960389.
16. Cho, Seongjae & Kang, In & Kamins, Theodore & Park, Byung-Gook & Harris, James. (2011). “Silicon-compatible compound semiconductor tunnelling field-effect transistor for high performance and low standby power operation”. *Applied Physics Letters.* 99. DOI: 10.1063/1.3670325.
17. Cho, Seongjae & Kim, Kyung & Park, Byung-Gook & Kang, In. (2011). “RF Performance and Small-Signal Parameter Extraction of Junctionless Silicon Nanowire MOSFETs”. *IEEE Transactions on Electron Devices*, vol. 58, pp. 1388 – 1396, DOI: 10.1109/TED.2011.2109724.
18. A. Kranti and G. A. Armstrong, “Design and optimization of FinFETs for ultra-low-voltage analog applications,” *IEEE Trans. Electron Devices*, vol. 54, no. 12, pp. 3308–3316, 2007, DOI: 10.1109/TED.2007.908596.
19. Madan, Himanshu & Saripalli, Vinay & Liu, Huichu & Datta, S.. (2012). “Asymmetric Tunnel Field-Effect Transistors as Frequency Multipliers”. *IEEE Electron Device Letters*. Vol. 33, pp. 1547-1549, 2012, DOI: 10.1109/LED.2012.2214201.
20. M. Berg *et al.*, “Imaging Atomically Thin Semiconductors Beneath Dielectrics via Deep Ultraviolet Photoemission Electron Microscopy,” *Phys. Rev. Appl.*, vol. 12, no. 6, p. 1, 2019.
21. A. J. L. Hofmann *et al.*, “Dipolar Doping of Organic Semiconductors to Enhance Carrier Injection,” *Phys. Rev. Appl.*, vol. 12, no. 6, p. 1, 2019.
22. S. Gupta and A. Nandi, “Effect of air spacer on analog performance of underlap tri-gate FinFET,” *Superlattices Microstructures.*, vol. 109, pp. 693–701, 2017, DOI: 10.1016/j.spmi.2017.05.053.

View publication stats

2021 2nd International Conference for Emerging Technology (INCET)
Belgaum, India. May 21-23, 2021

CFOA Based Negative Floating Fractional Order Capacitance Multiplier

Rahul Gupta
Department of Electronics and Communication
Delhi Technological University
Delhi, India
rahul.gupta9246@yahoo.in

Charvi Agrawal
Department of Electronics and Communication
Delhi Technological University
Delhi, India
charviagrawal4@gmail.com

Ankit Gupta
Department of Electronics and Communication
Delhi Technological University
Delhi, India
ayangupta137@gmail.com

Rajeshwari Pandey
Department of Electronics and Communication
Delhi Technological University
Delhi, India
rpandey@dce.ac.in

2021 2nd International Conference for Emerging Technology (INCET) | 978-1-7281-7029-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/INCET51464.2021.9456355

***Abstract*- In this paper, a Negative Floating Fractional Order Capacitance Multiplier has been presented using a fractional-order capacitor and Current Feedback Operational Amplifier (CFOA). The paper depicts the working of the proposed design through PSPICE simulations. Simulations of three instances of Negative Capacitor are presented.**

The proposed designs have been verified using a commercially available Current Feedback Op-amp macro model (IC AD-844A) and a fractional capacitance (FC) using RC Domino Ladder Network. The simulation results closely verify our theoretical propositions.

Keywords- CFOA, Capacitance, Fractance, Multiplier, Fractional-order, RC-circuit, Impedance, Phase

I. INTRODUCTION

A negative capacitance is synthetic impedance which can be created using an active block and few passive circuit components. The negative capacitances find wide applications in electrical/electronic signal processing such as to cancel the parasitic capacitance of inductors, transformers and heat-sinks [1]. The successful applications of these cancellations in the circuits can greatly benefit EMI noise suppression. The concept of negative capacitance can be used for bandwidth enhancement of a circuit by cancelling the speed bottleneck capacitance. The negative capacitance shunts can be used to reduce structural vibrations of flexural systems [2].

It is well known that realizing a physically large capacitor in integrated circuits (ICs) is a very costly affair and involves complex fabrication process steps. In many applications the physical size of the capacitor needed presents a major challenge such as CMOS image sensor which requires large noise attenuating capacitor and monolithic phase locked loop filter. A capacitance multiplier circuit is an alternate solution to realize large capacitances and is based on the principle of tunable impedance converter. The capacitance multiplier circuits are built around active blocks and small physical capacitance. A variety of capacitance multiplier circuits are available in literature [3-6] and references cited their in. The concept of capacitance multiplication can be extended to negative capacitance circuits to obtain the desired value of negative capacitance.

It has been established through research works proposed in past two decades that fractional calculus [7] acts as a more efficient mathematical tool for modeling real-world objects and processes than its integral counter-part. The fractional- order (FO) circuits and systems are widely being used in the field of signal processing, control systems, biomedical instrumentation, and many more.

The aim of this paper is to generalize the narrow integer-order negative capacitance multiplier circuits to more general fractional-order counterparts having flexibility of varying its fractional order. Thus, this paper presents a fractional order negative capacitance multiplier using two CFOA, a fractional capacitor [8] and two resistances.

The CFOA [9] is an active element that offers relatively wider range of operation frequency and higher value of slew-rate as compared to conventional op-amps owing to the use of current as the feedback signal.

Rest of the paper is structured as below: Section II highlights the design methodology of proposed negative fractional order floating capacitance multiplier. The simulation results are compiled in Section III to draw constructive inferences. The application of the proposed circuit is shown in Section IV and the conclusions have been presented in section V.

II. PROPOSED CONFIGURATION

A fractional order Laplacian operator is used to represent FO circuits. Using Laplace transform the fractional order operator can be simplified into integer approximation and thus it becomes physically realizable with fractional order elements (FOE).

The impedance function of a simple fractional-order element is given as,

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:51:30 UTC from IEEE Xplore. Restrictions apply.

$$Z(s) = (s^{\alpha}F) \quad (1)$$

where F is known as the fractance and α (a fractional number) is known as the exponent of the FOE. A positive α corresponds to a fractional inductor (FI) while negative fraction corresponds to fractional capacitor (FC). The impedance and phase angle of an FC are expressed respectively as [8]:

$$|Z(j\omega)| = 1/(\omega^{\alpha}F) \quad (2.a)$$

$$\theta = -\alpha\pi/2 \quad (2.b)$$

As of now the FCs are not commercially available however, can be emulated based on various structures [10] such as passive resistor-capacitor (RC) elements arranged in the form of RC tree, cross RC ladder, domino ladder etc. obtained using various rational approximations. In this work a truncated RC (resistor–capacitor) domino ladder-type network based on continued fraction Expansion (CFE) is used. The RC ladder network and the symbol for FC are shown Fig.1.

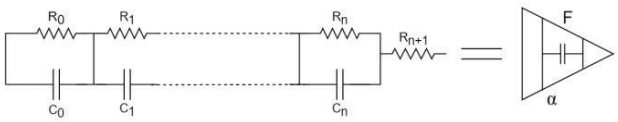


Fig. 1. RC Domino Ladder for Fractance and its corresponding symbol [8]

The symbolic diagram of CFOA is shown in Fig. 2. It consists of four ports. Considering the direction of port currents towards the CFOA block the port relationship, in matrix form [11] is given as:

$$\begin{bmatrix} I_y \\ V_x \\ I_z \\ V_w \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} V_y \\ I_x \\ V_z \end{bmatrix}$$

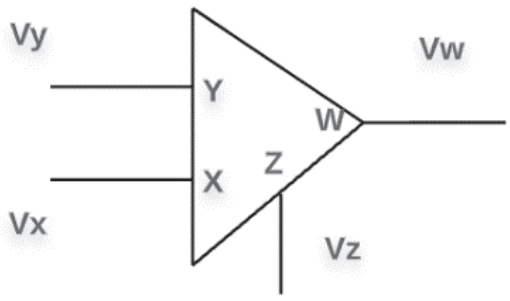


Fig. 2. Block Diagram of CFOA

The proposed Negative Floating Fractional Capacitance Multiplier is shown in fig. 3.

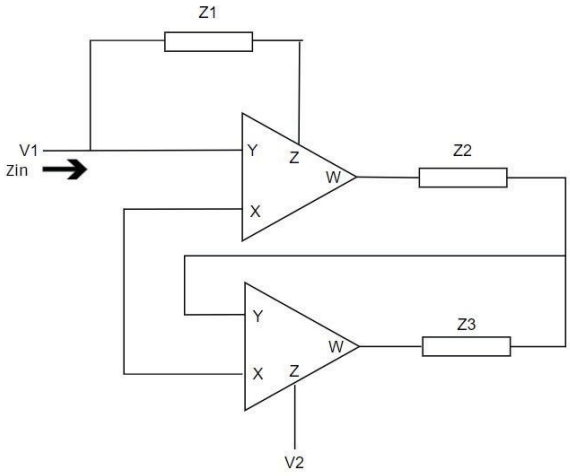


Fig. 3. Negative Floating Fractional Capacitance Multiplier

Expression for input impedance Z_{in}, is expressed as (3) [5]:

$$\frac{V_1 - V_2}{I_{in}} = -\frac{Z_1 Z_3}{Z_2} \quad (3)$$

From Equation (3), it can be observed that two configurations of the circuit are possible which are shown below:

Configuration 1: Selecting Z_1 as FC and Z_2 and Z_3 are resistors as shown in Fig.3 the impedance of the structure can be expressed as:

$$Z_{in} = -\frac{R_3}{s^{\alpha}FR_2} \quad (4)$$

Thus, Z_{in} is representing a Negative Fractional-order Capacitance of value,

$$C_{eq} = -K_{m1}s^{\alpha}F \quad (5)$$

Here K_{m1} denotes the multiplication factor whose mathematical representation is given in (5)

$$K_{m1} = \frac{R_2}{R_3} \quad (6)$$

The magnitude and phase angle of the proposed circuit are expressed as (6):

$$|C_{eq}(j\omega)| = K_{m1}(\omega)^{\alpha}F \quad (7.a)$$

$$\theta = \pi - \alpha\pi/2 \quad (7.b)$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:51:30 UTC from IEEE Xplore. Restrictions apply.

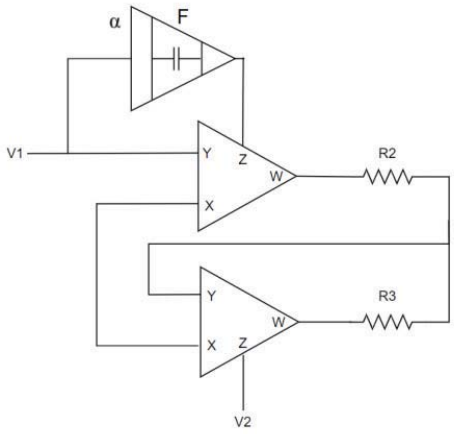


Fig. 4. Proposed configuration 1 of Negative Floating Fractional Capacitance

Configuration 2: If Z_3 is taken as FC and Z_1 and Z_2 are resistors as shown in fig. 4 then the impedance can be calculated as:

$$Z_{in} = -\frac{R_1}{s^{\alpha} F R_2} \tag{8}$$

thus, Z_{in} is representing a Negative Fractional-order Capacitance of value,

$$C_{eq} = -K_{m2} s^{\alpha} F \tag{9}$$

Here K_{m2} denotes the multiplication factor whose mathematical representation is given in (10)

$$K_{m2} = \frac{R_2}{R_1} \tag{10}$$

The magnitude and phase angle of the proposed circuit are expressed as (11):

$$\left|C_{eq}(j\omega)\right| = K_{m2}(\omega)^{\alpha} F \tag{11.a}$$

$$\theta = \pi - \alpha\pi/2 \tag{11.b}$$

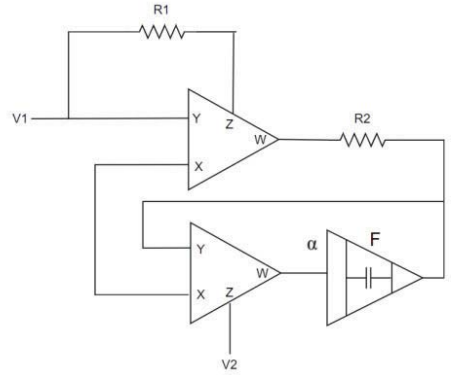


Fig. 5. Proposed configuration 2 of Negative Floating Fractional Capacitance

Thus, the magnitude and the order of negative FC can be varied by varying the multiplication factor K_{m1} or K_{m2} (as the case may be) the value of α. Thus, two degrees of freedom are possible in this configuration.

III. Simulation Results

The functionality of the proposed circuit is verified through SPICE simulations using AD844A macro model of CFOA with ±5 V supply voltages. Simulations are performed for two FC instances (i) F=12.5μF, α=0.3; (ii) F=3.75 μF, α=0.5 for which the components are listed in Table 1.

For F=12.5μF and α=0.3, the impedance magnitude is plotted for different values of K_m in Fig. 6 and it can be observed that the impedance decreases with the increase in frequency which confirms the expression for capacitance relation. The phase is plotted for the same in fig. 7 and yields a nearly constant phase which closely resemble the theoretical value of 153° for values of K_m 1-10 and starts deviating for larger K_m. Lastly for the same F and α, the transient response for phase relation between voltage and current is shown in fig. 8.

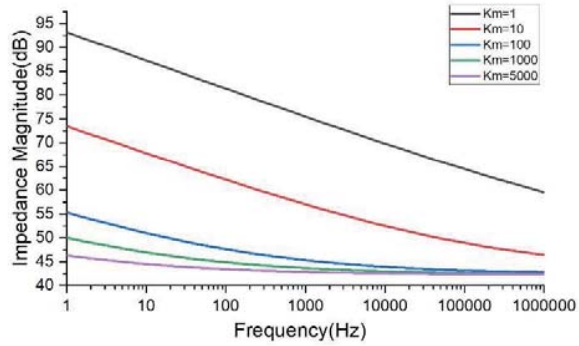


Fig. 6. Impedance Magnitude Response for **α**=0.3 and F=12.5u

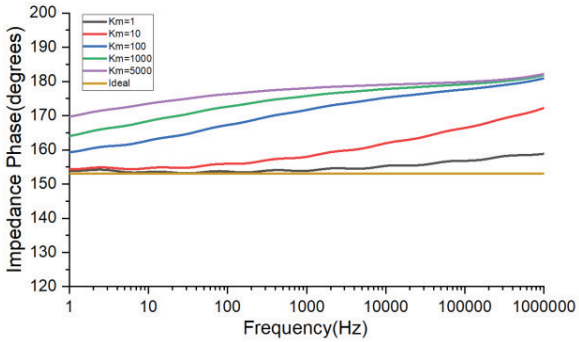


Fig. 7. Phase Magnitude Response for **α**=0.3 and F=12.5u

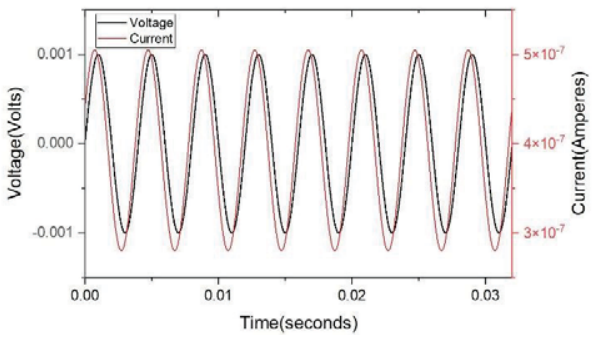


Fig. 8. Transient Response of Negative Capacitance with α=0.3. F=12.5u and K_m =1

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:51:30 UTC from IEEE Xplore. Restrictions apply.

For F=3.75u and α=0.5, the impedance magnitude is plotted for different values of K_m in fig. 9 and it can be observed that the impedance decreases with the increase in frequency which confirms the expression for capacitance relation. The phase is plotted for the same in fig. 10 and yields a nearly constant phase which closely resemble the theoretical value of 135° for values of K_m 1-10 and starts deviating for larger K_m. Lastly for the same F and α, transient response for phase relation between voltage and current is shown in fig. 11.

By examining the magnitude and phase plot of the input impedance of all the different instances, the variability of the multiplication factor can be defined up to 1000. Some deviations from ideal plot are observed, which are majorly due to internal parasitic elements of CFOA, thus limiting the operational bandwidth. Therefore, the multiplier circuit has a practical frequency range from 1 Hz to 10 KHz.

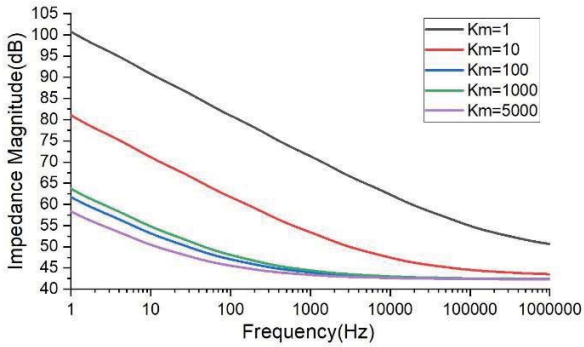


Fig. 9. Impedance Magnitude Response for α=0.5 and F=3.75u

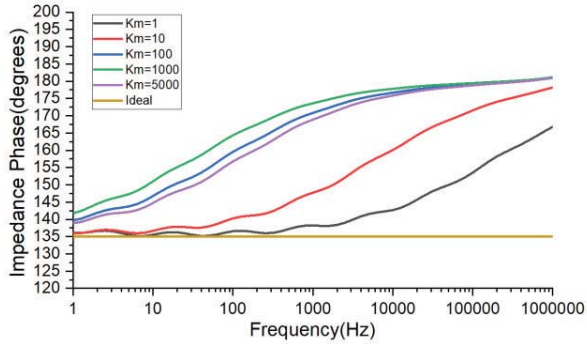


Fig. 10. Phase Magnitude Response for α=0.5. and F=3.75u

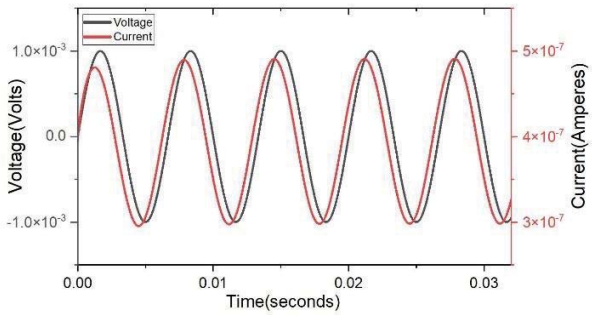


Fig. 11. Transient Response of Negative Capacitance with α=0.5. F=3.75u and K_m =1

IV. APPLICATION

During circuit designing often stray capacitances are generated which can hamper the performance. These capacitances need not be of integral order, therefore in such case a negative fractional capacitance multiplier can be used to design cancellation circuits to get rid of such problems. Fig. 12 below shows an accurate description of this application. The circuit consists of a resistor in series with a parallel combination of a parasitic positive fractional capacitor C_1 and proposed negative fractional capacitor C_2, which can be easily tuned to the exact value in order to cancel the effect of C_1 using the proposed design.

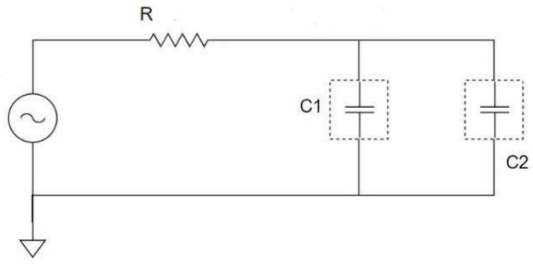


Fig. 12. Circuit for Capacitor Cancellation

The input impedance of the circuit in fig. 12 is expressed as (12)

$$Z_{in} = R + \frac{1}{s(C_1 - |C_2|)} \quad (12)$$

By properly tuning the domino capacitance using the proposed fractional multiplier circuit we can make the circuit purely resistive, $C_1 = |C_2|$, reducing expression (12) to (13).

$$Z_{in} = R \quad (13)$$

TABLE I. VALUES OF ELEMENTS OF RC LADDER [2]

	Ladder elements	Block Numbers												
		0	1	2	3	4	5	6	7	8	9	10	11	12
F=12.5 μF α=0.3	R (Ω)	46.9k	21k	12.6k	7.61k	4.58k	2.76k	1.66k	999	601	362	218	65.6	259
	C (F)	22.2u	9.12u	2.79u	854n	261n	80n	24.5n	7.49n	2.29n	701p	215p	131p	
F=3.75 μF α=0.5	R(Ω)	336k	83.7k	32.5k	12.6k	4.89k	1.90k	736	286	111	43	8.34	17.6	
	C (F)	4.98u	3u	1.16u	452n	176n	68.1n	26.4n	10.3n	3.98n	1.54n	1.2n		
	C (F)	149.3n	151.2n	77.1n	39.3n	20.1n	10.2n	5.2n	2.66n	1.36n	1.38n			

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:51:30 UTC from IEEE Xplore. Restrictions apply.

The simulation of the cancellation of capacitors as proposed in fig. 12 is shown in fig. 13 wherein the current and voltage are in phase as the effect of the fractional capacitor is canceled by the negative fractional capacitor tuned to the appropriate value thereby making the circuit purely resistive.

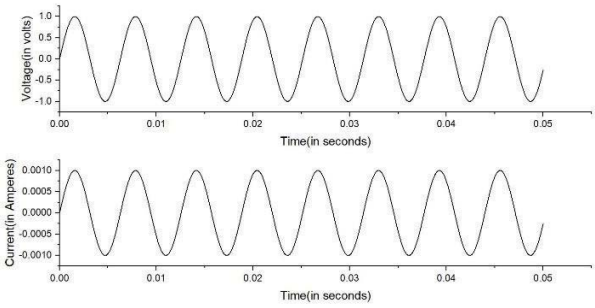


Fig. 13. Transient Response showing cancellation

V. CONCLUSION

In this paper, a negative floating fractional capacitance multiplier designed using CFOA has been presented. The proposed design is verified with the use of SPICE simulations for fractional-order of 0.3 and 0.5. The behavior of the proposed design was tested by comparing the theoretical impedance magnitude and phase responses for different values of the multiplication factor with the ideal ones.

The proposed capacitance multiplier shows promising results in generating high value fractional capacitances with minimum hardware requirement and also showed successful outcomes in cancellation circuit applications.

REFERENCES

[1] S. Wang and F. C. Lee, "Negative Capacitance and its Applications on Parasitic Cancellation for EMI Noise Suppression," *2007 IEEE Power Electronics Specialists Conference*, Orlando, FL, 2007, pp. 2887-2891

[2] Benjamin S Beck, Kenneth A Cunefare and Manuel Collet, "The power output and efficiency of a negative capacitance shu nt for vibration control of a flexu ral system", Published 30 April 2013, 2013 IOP Publishing Ltd

[3] S. Unhavanich, O. Onajn, and W. Tangsrirat “Tunable Capacitance Multiplier with a Single Voltage Differencing Buffered Amplifier,” In Proc. of the International MultiConference of Engineers and Computer Scientists (IMECS), Hong Kong, China, 2016, pp. 1–4.

[4] W. Tangsrirat, “Resistorless tunable capacitance multiplier using single voltage differencing inverting buffered amplifier,” Rev. Roum. Sci. Techn.– Électrotechn. et Énerg., vol. 62, no. 1, pp. 72–75, 2017.

[5] W. Tangsrirat, and O. Channumsin, “Tunable Floating Capacitance Multiplier Using Single Fully Balanced Voltage Differencing Buffered Amplifier,” J. Commun. Technol. Electron, vol. 64, pp. 797–803, 2019.

[6] M. A. Al-Absi and A. A. Al-Khulaifi, “A New Floating and Tunable Capacitance Multiplier with Large Multiplication Factor,” IEEE Access, vol. 7, pp. 120076–120081, 2019.

[7] A. G. Radwan and K. N. Salama, "Passive and Active Elements Using Fractional $L_\alpha C_\beta$ Circuit," IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 58, no. 10, pp. 2388-2397, Oct. 2011.

[8] A. Adhikary, P. Sen, S. Sen and K. Biswas, “Design and Performance Study of Dynamic Fractors in Any of the Four Quadrants,” in Circuits, Systems, and Signal Processing , vol. 35, no. 6, pp. 1909–1932, Jun. 2016.

[9] F.J. Lidgey, K. Hayatleh, 'Current-feedback operational amplifiers and applications', Electronics & Communication Engineering Journal, 1997, 9, (4), p. 176-182.

[10] R. Verma, N. Pandey, & R. Pandey, "Realization of a higher fractional order element based on novel OTA based IIMC and its application in filter" in Analog Integr Circ Sig Process 97, 177–191 (2018).

[11] S. Singh, N. Pandey, R. Pandey. (2020) CFOA Based Negative Floating Capacitance Multiplier. In: Zakaria Z., Ahmad R. (eds) Advances in Electronics Engineering. Lecture Notes in Electrical Engineering, vol 619. Springer, Singapore

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:51:30 UTC from IEEE Xplore. Restrictions apply.

Classification of ancient inscription images on the basis of material of the inscriptions

N. Jayanthi
Department of ECE
Delhi Technological University
Delhi, India
njayanthi@dce.ac.in

Tarush Sharma
Department of ECE
Delhi Technological University
Delhi, India
tarushsharma_2k17ec173@dtu.ac.in

Vinay Sharma
Department of ECE
Delhi Technological University
Delhi, India
vinaysharma_2k17ec185@dtu.ac.in

Shubham Tyagi
Department of ECE
Delhi Technological University
Delhi, India
shubham_2k17ec155@dtu.ac.in

S. Indu
Department of ECE
Delhi Technological University
Delhi, India
s.indu@dtu.ac.in

2021 3rd International Conference on Signal Processing and Communication (ICPSC) | 978-1-6654-2864-4/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICSPC51351.2021.9451641

***Abstract*—Machine Learning and AI has allowed us to process images and help us solve vital problems. In this paper, we are classifying inscription images into three image inscription classes namely stone inscriptions, metal inscriptions and palm leaves inscriptions. Due to decaying materials of ancient inscriptions, the classification of such materials becomes challenging. To address this problem, we are using various feature extraction methods like GLCM, KAZE, BRISK to implement texture based feature detection and subsequently classifying them. Both linear and non-linear methods for feature extraction are being used. The paper is concluded by performing classification of images and also making comparisons between all the feature extraction methods in terms of memory required by the algorithm, time scalability and accuracy score of the method. The paper consists of rich information which would be useful in decision making in classification problems.**

***Index Terms*—GLCM, KAZE, BRISK, ancient inscriptions, texture classification**

I. Introduction

One of the integral part in the field of the image processing is the classification based on texture of image.
Images are stored in the form of mathematical data representing height and width of the images with depth of the pixels representing the color of the image.
Most meaningful features for describing images are very much the same features which human being uses in order to interpret the images. Three fundamental pattern element used by humans to interpret colour images are spectral, textural and contextual features. Average variation in tone in various visible and infrared portions of an electromagnetic spectrum is described by spectral features, whereas the spatial distribution in the variations of tone in a band is described by textural features. Contextual features from around the area under observation, provides valuable insights from the pictorial data blocks. If the images being processed are black and white, then tone and the texture of the image are the most important features.

Texture can be described as the fineness, coarseness or the smoothness of an image. It also tells if the image is molled, irregular, lineated etc. Texture is a property possessed by all the surfaces, be it a piece of wood, a grain of sand or a piece of silk. Textures contains in it the information of the structural arrangement of the surfaces and their relationship with the surroundings.

II. Related Works

Texture is a fundamental characteristic used in identifying regions of interest in an image, extracting features from images is a well researched topic and various models are proposed.

GLCM which stands for Gray Level Co-occurrence Matrix,a feature extraction method known for very long was proposed by Haralick.Since it's inception, GLCM has been one of the most important feature extraction methods and finds use in various application of texture analysis.Fourteen different feature were from GLCM were identified and proposed by Haralick, to specify texture [1].Another model which implements feature extraction using GLCM and classification using CBIR(content based image retrieval) is proposed in [2].

Texture classification can be done in many different ways,a dominant neighbourhood structure is proposed in [3], which incorporates the global image features for this purpose.Invariant features of local textures(ILFT) is proposed in [4] which is a texture descriptor algorithm.

S. Leutenegger et al. introduced BRISK method in 2011 [5], which which uses AGAST algorithm to determine the presence of corners, then incorporating FAST corner score to filter the corners, while working simultaneously in the scale space pyramid searching for a maxima.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:28:58 UTC from IEEE Xplore. Restrictions apply.

Pablo Fernandez Alcantarilla et al. intorduced KAZE features, a 2D feature detection algorithm that exploits nonlinear scale spaces. [6].

III. Methodology

Inscription images are being classifies into three classes which are stone, metal and palm leaves inscription. To achieve this we performed the following steps.

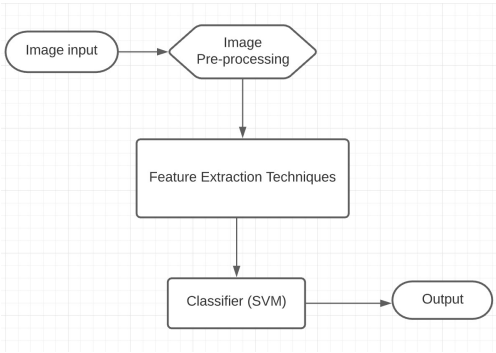


Fig. 1. Flowchart of proposed model

A. *Input & pre-processing*

The data-set consisting of stone, metal and palm-leaves inscription images was prepared from various sources. Due to decaying materials and bad resolution of images classification of data-set is quite difficult. To achieve desired results various pre-processing steps like size reduction of images to appropriate size and in some feature extraction method like GLCM we are converting images into gray-scale. This data is then fed into the various feature extraction algorithms.

B. *Feature Extraction techniques*

We are using various feature extraction methods to extract feature vectors from the images which are essential for the classifier to classifying the images. The feature extraction methods given below like GLCM with its variations, KAZE and BRISK are being used.

1. GLCM

To identify region of interest in an image, texture is a key characteristic. GLCM is used a feature extraction method, where an image can be represented in a matrix that can detect special nature of gray scale pixel distribution in image texture. It evaluates the spatial relationship between pixels, by determining how frequently a pair of pixels with a specific value occur in the image, and thereby creating a matrix.

These are 14 haralick features derived from GLCM matrix:

1. Maximal Correlation Coefficient

$\mathrm{f}_1 = (secondlargesteigenvalueofQ)^{1/2}$
$where,$
$Q(i,j) = \sum_k \frac{p(i,k)p(j,k)}{p_x(i)p_y(k)}$

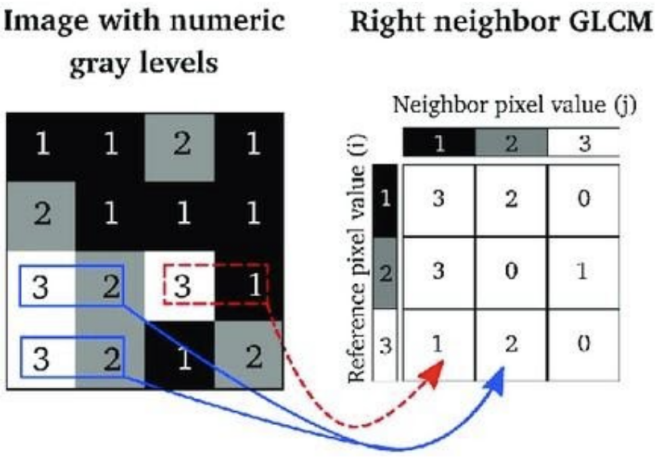


Fig. 2. GLCM matrix

2. Information Measure of Correlation 1

$\mathrm{f}_2 = \frac{HXY - HXY1}{max(HX,HY)}$

3. Difference Variance

$\mathrm{f}_3 = variance.of.p_{x-y}$

4. Sum Entropy

$\mathrm{f}_4 = -\sum_{i=2}^{2N_g} p_{x+y}(i) log\left\{p_{x+y}(i)\right\}$

5. Sum Average

$\mathrm{f}_5 = \sum_{i=2}^{2N_g} i p_{x+y}(i)$

6. Sum of Squares: Variance

$\mathrm{f}_6 = \sum_i \sum_j (i-\mu)^2 p(i,j)$

7. Contrast

$\mathrm{f}_7 = \sum_{n=0}^{N_g-1} n^2 \left\{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)\right\}$

8. Angular Second Moment

$\mathrm{f}_8 = \sum_{i=1}\sum_{j=1} p(i,j)^2$

9. Correlation

$\mathrm{f}_9 = \frac{\sum_i \sum_j (ij)p(i,j) - \mu_x\mu_y}{\sigma_x\sigma_y}$

10. Inverse Difference Moment

$\mathrm{f}_{10} = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:28:58 UTC from IEEE Xplore. Restrictions apply.

11. Sum Variance

$$f_{11} = \sum_{i=2}^{2N_g} (i - f_4)^2 p_{x+y}(i)$$

12. Entropy

$$f_{12} = -\sum_i \sum_j p(i,j) log(p(i,j))$$

13. Difference Entropy

$$f_{13} = -\sum_{i=0}^{N_g-1} p_{x-y}(i) log\{p_{x-y}(i)\}$$

14. Information Measure of Correlation 2

$$f_{14} = (1 - exp[-2.0(HXY2 - HXY)])^{1/2}$$

2. KAZE

KAZE feature extraction method is a novel method which works in a non-linear scale on 2d feature detection and description. To identify features in nonlinear scale spaces, non linear diffusion is used which removes the noise and keep all the valid details right in place from the image. One of the simpler cases of diffusion in non-linear manner, variable conductance diffusion is being used. By using Additive Operator Splitting (AOS) schemes, an efficient non-linear space is being built.

A novel mathematical framework called Fast Explicit Diffusion (FED) is used in Accelerated-KAZE Features. Also, a robust Modified-Local Difference Binary (M-LDB) descriptor is computed that uses gradient inputs from non-linear scale space.

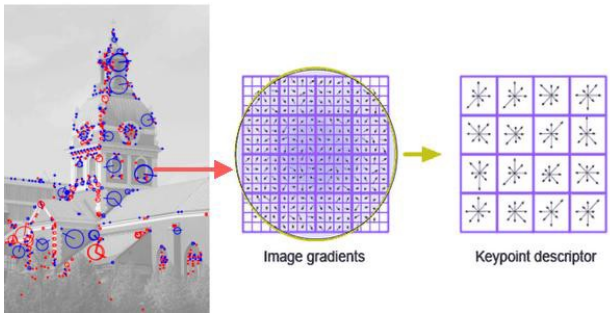


Fig. 3. KAZE

3. BRISK

Three main modules of BRISK algorithm are: descriptor matching, keypoint description and keypoint detection. Initially, the scale space pyramid is being computed, and the extreme points which are stable in continuous scale space of sub-pixel precision are extracted by AGAST (the Adaptive corner detection operator). From the local image neighbourhood, binary feature descriptor of the local image is made ,by incorporating the gray-scale relationship.

Finally, for the purpose of feature matching Hamming distance is used. For the purpose of realisation of scale invariance of descriptors and rotational invariance of descriptors,BRISK algorithm is used which in a multi-scale layer detects feature points and with the help of pixel pairs at along distance, tells the direction of master mode for realising both scale invariance and rotational invariance respectively.

BRISK is a slow method and requires a significant amount of memory for processing.

C. *Classifier*

A supervised machine learning model Support Vector Machine(SVM) was used which is the best fit for group classification problems. By feeding a set of labelled training data to the SVM model, the trained model then was easily able to categorize new input images. The main inclination to use SVM as classifier was because it performs very efficiently on a smaller data-set.

IV. Implementation

A. *Data-set*

We collected a data-set of around 150 images of stone inscriptions, palm leaf inscriptions and metal inscriptions. Data-set was collected from different sources. The images in the data-set were processed to fit our use-case on which we can train and run our model. There are limited algorithms that perform better on small data-sets.



Fig. 4. Stone Inscription

B. *Approach*

We have taken our data-set and trained all the mentioned model to make comparative analysis on all the methods. First we have performed classification using just SVM without using any feature extraction method so we can demonstrate the significance of using feature extraction method. But SVM as a classifier is being used in all the approaches.

1. SVM with kernel = RBF

Here, firstly the images were resized to 256 x 256 pixels and then converted to gray scale and fed to the SVM classifier. The accuracy of the model was found to be 64The accuracy

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:28:58 UTC from IEEE Xplore. Restrictions apply.



Fig. 5. Metal Inscription

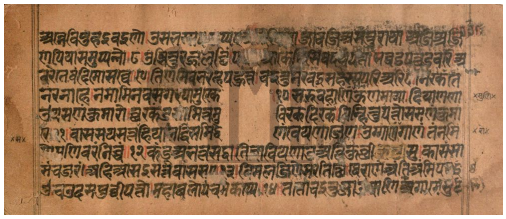


Fig. 6. Palm Inscription

wasn't up to the mark so we decided to look for other approaches.

2. GCLM with gray scale images with 14 features
Here, we fed gray scale images directly into GLCM without resizing them. We decided to use GLCM because it helps in feature extraction and texture based detection. We used haralick texture features which are derived from GLCM matrix and it is a set of 14 different features. For the classification purpose we used an SVM classifier. The output haralick texture features were fed directly to the SVM classifier.

3. GCLM with colored scale images with 14 features
The accuracy of the model was further improved by using coloured images as input instead of gray scale images. Coloured images are 3 dimensional hence 13 GLCM matrices were created which gave better feature extraction property which in turn gave better accuracy, while for gray scale images only 4 GLCM matrices are created per image which resulted in lower accuracy.

4. GCLM with gray scale images with 8 features
Here, we fed gray scale images directly into GLCM without resizing them. We decided to use GLCM because it helps in feature extraction and texture based detection. We used haralick texture features which are derived from GLCM matrix and it is a set of 14 different features. For the classification purpose we used an SVM classifier. The output haralick texture features were fed directly to the SVM classifier. The 8 feature used to train the model were - Sum of Squares: Variance, Difference Entropy, Maximal Correlation Coefficient, Contrast, Difference Variance, Information Measure of Correlation 1, Inverse Difference Moment and Information Measure of Correlation 2.

5. KAZE
The colored images are directly passed as input into KAZE extraction algorithm. There are a number of 2d feature points from which some are selected using a moving window .First, the program uses extracted features to detect key points on the image (the center point of the local pattern). From image to image key point number varies therefore to ensure the resulting feature vector size to remain same we need to add some rules (this is because we can't compare vectors with different dimensions when calculating).Thereafter we compute vector descriptors on the basis of key points. The size of each descriptor is 64. We have 32 such descriptors, so the dimensions of our feature vector is 2048. The output feature vector was directly fed into SVM for classification.

6. BRISK
The colored images are directly passed as input into BRISK extraction algorithm. Here also, similar to KAZE algorithm procedure feature extraction is carried out. BRISK is a very time consuming method. There are a number of 2d feature points used to detect feature in this method, a feature vector of 2048 features is used from which output features are calculated. The output feature vector was directly fed into the SVM for classification.

V. Results & Discussions

A. SVM with kernel = RBF

For a smaller data-set just SVM does not give us an appreciable accuracy. The accuracy reaches upto 64.3%.

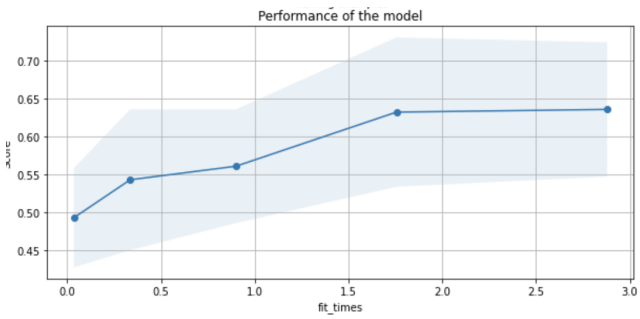


Fig. 7. Performance evaluation for SVM

B. GLCM with gray scale images with 14 features

Gray scale images with all 14 features helps in improving the accuracy. The accuracy goes upto 71.2%.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:28:58 UTC from IEEE Xplore. Restrictions apply.

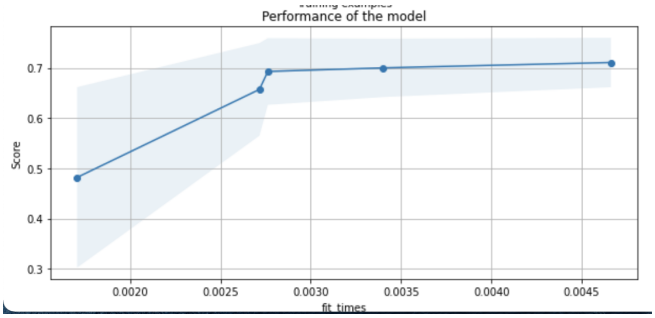


Fig. 8. Performance evaluation for GCLM(gray-scale)

C. GLCM with colored scale images with 14 features

Colored images with all 14 features further improves the accuracy. The accuracy goes up to 73.3%.

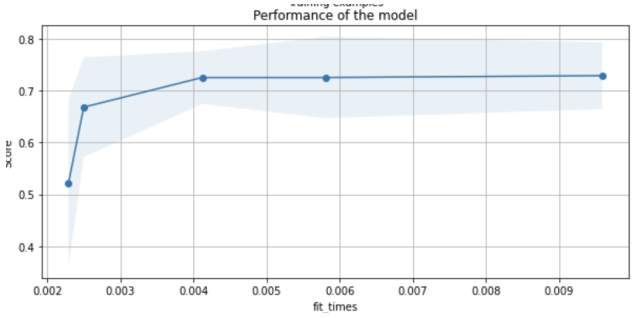


Fig. 9. Performance evaluation for GLCM(colored)

D. GLCM with gray scale images with 8 features

Gray scale images with only 8 features after removing the redundant features from our model. The accuracy goes up to 74.56%.

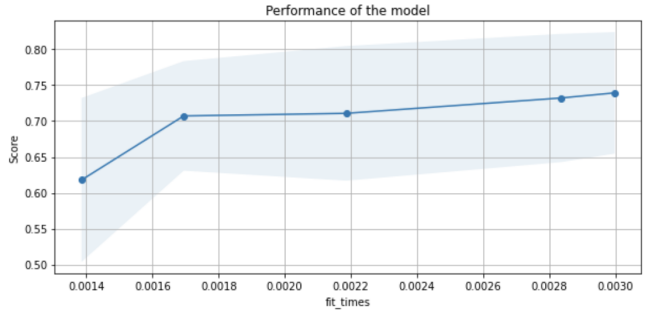


Fig. 10. Performance evaluation for GLCM(selective features)

E. KAZE

KAZE feature extraction method with 1024 feature points was used. The accuracy goes up to 80.5%.

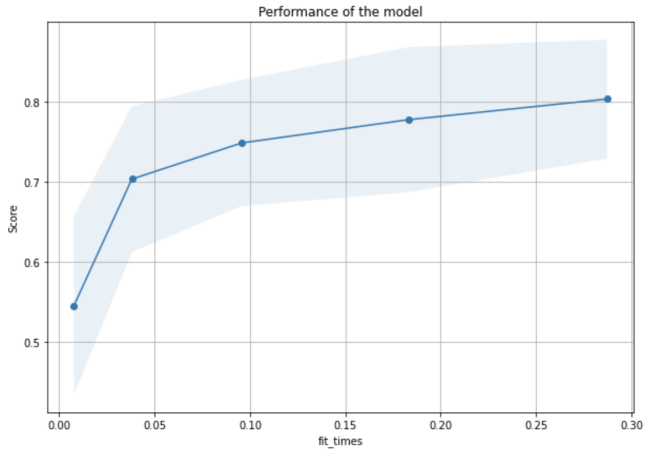


Fig. 11. Performance evaluation for KAZE

F. BRISK

BRISK feature extraction method with 1024 feature points was used. The accuracy goes up to 65%.

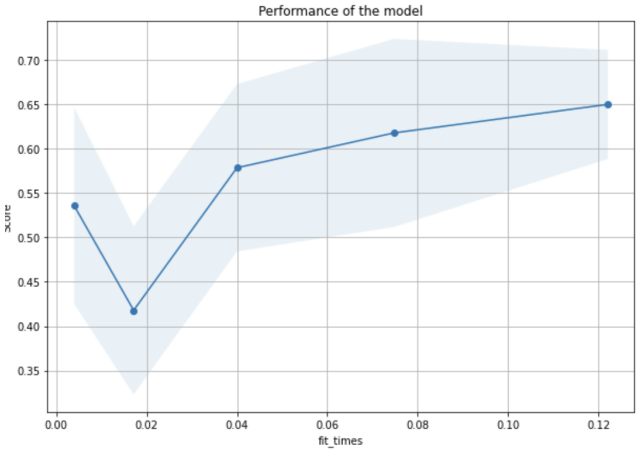


Fig. 12. Performance evaluation for BRISK

TABLE I
ACCURACY SCORES OF EACH APPROACH

Approach	*Accuracy*
SVM(kernel = RBF)	64.3
BRISK (2048 feature points)	65
GLCM(grayscale images & 14 features)	71.2
GLCM(coloured images & 14 features)	73.3
GLCM(grayscale images & 8 features)	74.56
KAZE (2048 feature points)	80.5

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:28:58 UTC from IEEE Xplore. Restrictions apply.

Here, exhaustive comparison between GLCM and its variations, BRISK and KAZE is being done. The experimental results presents rich insights which help in making critical conclusions in feature extraction and classification methods.

SVM model being the elementary model has least accuracy. Because it does not consider any features in the image it just classifies in the basis of pixel values.

GLCM considers spatial relationship of pixels hence it was a better fit for classifying the images. A better classification accuracy score was achieved in comparison with SVM. Also, GLCM is computationally most efficient and requires least memory in comparison with other methods. Gray scale images are represented in 2d matrix therefore only 4 GLCM matrices can be obtained for feature extraction but using colored images as input which are represented in 3d matrices from which 13 GLCM matrices can be obtained which assists in better feature extraction hence better classification. Further, there were various features which were having similar values for the three classes of inscriptions (stone, metal and palm leaves) hence, removing the redundant features further improved the accuracy score.

BRISK uses Hamming distance for feature matching which comparatively results in an inefficient feature extraction. Therefore, this results in a poor classification accuracy.

KAZE outperforms all the above feature extraction methods when it comes to classification of images into three classes. KAZE identifies features in a nonlinear scale space using non linear diffusion which removes the noise and keep all the valid details right in place from the image. Therefore, it results in the highest accuracy score.

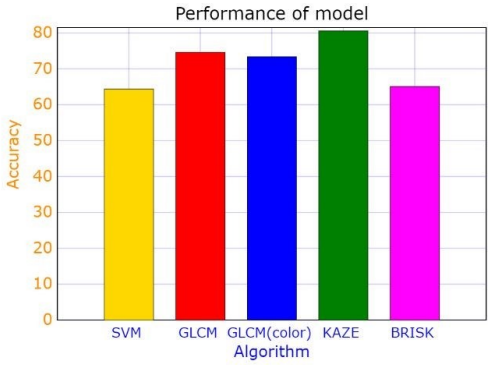


Fig. 13. Performance of model

VI. CONCLUSION & FUTURE WORK

In this paper, we are performing classification of inscription images based on the materials used. Different feature extraction methods are being used to extract feature matrices which are subsequently being used to perform classification. Accuracy of classification increases substantially because of feature extraction method. Further, we are also evaluating and making comparative analysis between different kinds of feature extraction methods. Comparative analysis shows that KAZE feature algorithm is the most efficient algorithm in terms of accuracy with accuracy score of 80.5%. While KAZE is the most efficient method, taking time scalability and memory requirement into account GLCM turns out to be a better option. Alternatively, in place of feature extraction methods used in this paper neural networks like CNN can be used. But it requires a larger data-set which is one of the challenges and work for future in this field.

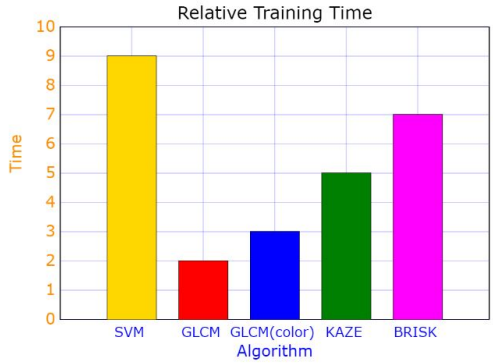


Fig. 14. Relative training time

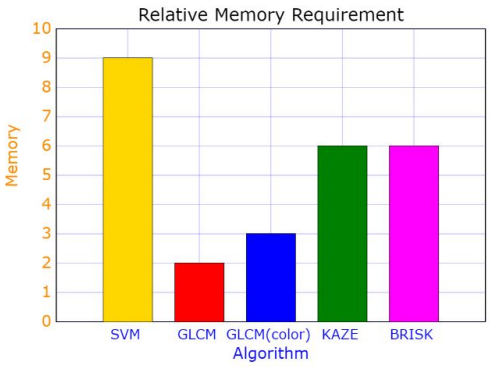


Fig. 15. Relative memory requirement

REFERENCES

[1] R. Haralick, K. Shanmugam, and I. Dinstein, (1973) "Textural Features for Image Classification", IEEE Trans. on Systems, Man and Cybernetics, SMC–3(6):610–621

[2] Yuslena Sari, Muhammad Alkaff, and Ricardus Anggi Pramunendar, " Classification of coastal and Inland Batik using GLCM and Canberra Distance ",AIP Conference Proceedings 1977, 020045 (2018)

[3] Fakhry M. Khellah, "Texture Classification Using Dominant Neighborhood Structure", IEEE Transactions on Image Processing, 2011, pp 3270-3279

[4] P. Janney and G. Geers, "Texture classification using invariant features of local textures", Image Processing, IET, 2010, pp 158-171.

[5] S. Leutenegger et al., "BRISK: Binary robust invariant scalable keypoints," in IEEE International Conference on Computer Vision, Barcelona, ICCV, 2011, pp. 2548-2555

[6] P. F. Alcantarilla et al., "KAZE features," in European Conference on Computer Vision, Berlin, ECCV, 2012, pp. 214-227

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:28:58 UTC from IEEE Xplore. Restrictions apply.

Classification of Human Activities using data captured through a smartphone using deep learning techniques

Lokesh Dhammi
Dept. of Electronics and Communication
Delhi Technological University(DTU)
Delhi, India
lokesh.dhammi@gmail.com

Piyush Tewari
Dept. of Electronics and Communication
Delhi Technological University(DTU)
Delhi, India
piyushtewari@dtu.ac.in

2021 3rd International Conference on Signal Processing and Communication (ICPSC) | 978-1-6654-2864-4/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICSPC51351.2021.9451772

***Abstract*—Classification and recognition of activities performed by humans is an essential task when we are into an era of connected-sensors and Industry 4.0, also referred to as the Internet of Things (IoT). In this paper, we propose an intelligent human activity-recognition system that can automatically predict daily human activities based on data acquired by sensors, using deep learning techniques like Two-dimensional Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM). The data collected using the sensors is fed to CNN and LSTM network. The proposed model is trained and tested on the WISDM dataset. The proposed deep learning architecture has Batch Normalization (BN) and dropout layers to avoid overfitting. The proposed architecture is of low complexity and achieves high accuracy and F1 score. The final accuracy obtained using the LSTM model is up to 97.5% with 21,922 trainable parameters. The obtained results are represented with the help of a convolutional matrix and learning curves. The results are quite promising and can be utilized without any restrictions on environments or domain-structures.**

Keywords—Human Activity Recognition (HAR), wearable sensors, deep learning, Convolutional Neural Network (CNN) Recurrent Neural Network (RNN), Long-Short-Term-Memory (LSTM)

I. INTRODUCTION

HAR aims to classify the input from the data received from the wearable sensors. In the era of IoT and AI, classifying human action has become a great demand for researchers and industrialists [1]. HAR has a vital role in the daily activities of humans due to its power to absorb the intense and far-reaching knowledge regarding the day-to-day actions performed by humans [2]. Hitherto, this futuristic technology of recognizing and predicting human actions is used in a variety of fields like behavior analysis [3], gesture recognition [4], GAIT analysis [5], and video surveillance [6], etc. A lot of astounding and stunning results are published and verified in the field of artificial technology (neural networks) and other technologies with an increased and improved system performance [7][8]. Technologies like deep learning, CNN and RNN have the ability to extract and learn the deep features, which is revolutionizing the world of computer vision and AI.

By virtue of advanced technology available for circuit optimization, embedded sensors can accurately sense and transmit information about human activities [7][8][9]. A variety of sensors are used according to their cost of installation, price and the type of information to be fetched (e.g., GPS, accelerometer, camera, etc) [10]. Sensors used for HAR applications can be fixed sensors or/and movable sensors [11]. Fixed sensors fetch the data while mounted on a particular position, which includes 'acoustic-sensors' [12], 'static-cameras' [14], 'radars' [13], etc. Some of the feature extraction methods are optical-flow and background-subtraction methods. Movable sensors are those which can change their position or location while collecting the data like body-worn sensors or smartphone sensors. These sensors can be a temperature sensor, speed sensor, object detection sensor, gyroscope sensor, accelerometer sensor, magnetometers, etc. Different parameters change due to the motion of the object under consideration, and can be fetched by the movable sensors. Rapid increase in the usage of sensors is observed due to miniaturization, flexibility, low cost, low power consumption and high accuracy. Today Smartphones with embedded sensors are quite capable to sense and capture the data related to any human activity. Many wearable sensors are also in the market which are to be worn by the individual on the body parts. These sensors are for elderlies, patients and children to monitor their activities for daily-living. Mishappenings can be avoided by monitoring their daily routine activities. But usually wearing the sensors on the body is very uncomfortable for patients. In such situations, the sensors which are already embedded in our smartphones can be extremely useful because they are very handy and portable. This will help the health care departments for better monitoring of patients.

Today all the gadgets are interconnected with each other and the smartphone is the central control unit. Smartphone plays a vital role in our daily life and has become an integral part of life. A smartphone can gather a variety and ample amount of information with the help of advanced software modules and sensors implanted in them. This practical use of smartphones will be of great help for recognizing the meaningful activities of the human body. The scientific research is continuously going on to classify near anatomical human actions based on data collected by the sensors available in smartphones [18].

Outline of this paper is as follows. In Section II we discuss the work related to the topic based on the review of literature. Section III depicts the data set used in our research work. Section IV explains the Methodology used to accomplish our research work. Section V and VI present the experimental results and conclusions obtained respectively.

II. INTRODUCTION TO RELATED WORK

HAR is a very recent and challenging area of research, and currently a lot of research is going on in this field. Initially, it started with learning and predicting simple actions not involving any moving object and varying backgrounds such as handshaking, clapping hands, etc. Research in this area is

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:33:43 UTC from IEEE Xplore. Restrictions apply.

gradually moving towards the prediction of complex real-life actions e.g., walking, kicking, diving, etc, by learning more and more features [15],[16].

Different kinds of sensors are used for extracting the features with the help of various technologies and innovations for recognizing human activities [18]-[25]. Initial research mainly focused on the traditional machine learning algorithms to classify the sensor-based data e.g. – Decision Tree Algorithm, Naive Base Algorithm, Support Vector Machine (SVM), etc [20]-[21]. L. Bao et al. and J. Mantyjarvi et al. presented their research work on "Motion sensors for classification". They attached the sensors on the body parts at multiple locations. L. Bao et al. [22] gathered data using the biaxial accelerometer positioned at four limbs and a hip. J. Mantyjarvi et al. [23] collected data using two triaxial accelerometers positioned at both sides of the waist. Tapia et al. [24] determined the action in the real-time by attaching the five-triaxial-accelerometer along with cardio-vascular monitor to measure the heart-rate. All multiple body-worn (wearable) sensor methods provide high accuracy. As mentioned in the introduction body-worn sensors are uncomfortable for users especially during their prolonged use and monitoring of the elderly-patients.

Machine learning techniques used for classification has shown positive outcomes. Jain et al. [25], have used 'Support-Vector-Machine (SVM) and K-Nearest-Neighbour (KNN)', methods on two public datasets to identify the actions. Jalloul et al. [26] applied Random-Forest (RF) algorithm and got an accuracy of 84.6% with the help of inertial measurements and separate monitoring setup. Margarito et al. [27] classified 8-activities using accelerometer-sensors worn on the wrist of the subject and applied template-matching to recognize the actions. Zhu et al. [28], proposed the Smart-Assisted Living (SAIL) system for disabled and elderly-patients. Bieber et al. [29] came forth with a user-friendly application that predicts day-to-day physical activities with the help of data gathered using sensors implanted in the smartphone.

Caros J S et al. [30], implemented the decision-tree algorithm to classify activities like 'sitting' and 'standing'. Maurer et al. [31] compared different algorithms for activity recognition and Wang S et al. [32] had used the Support-Vector-Machine (SVM) to recognize the activities. Machine-learning algorithms have given promising results but, they rely on heuristic manual-feature-extraction methods which are of finite knowledge especially in the case of HAR tasks.

To overcome this limitation, researchers have shifted their focus to the deep-learning techniques which automatically extract features from the sensor-gathered information at the time of training and present their enhanced version.

Deep-Learning has already shown very promising and successful results in areas like classification of images, voice-recognition, Natural Language Processing(NLP), and now also widely being used in the field of HAR [33]-[35]. Y. Zheng et al. [33] has collected the data using a triaxial (three-axis)-accelerometer and converted it into an 'image' and then applied CNN to learn features and classify the actions. In [34], Ordóñez and Roggen proposed an activity-recognition classifier with the combination of CNN network and LSTM(RNN) to distinguish different activities. The experimental outcomes were represented in terms of parameter- F1 score. In that study, CNN classifier got 0.93 F1 score and LSTM classifier got 0.958 F1score.

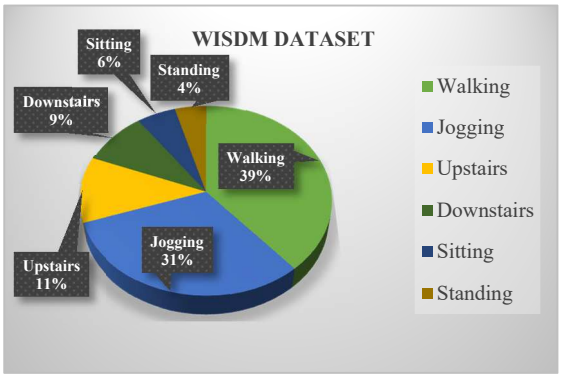


Fig. 1. A pie-chart showing the activities of WISDM Dataset [36]

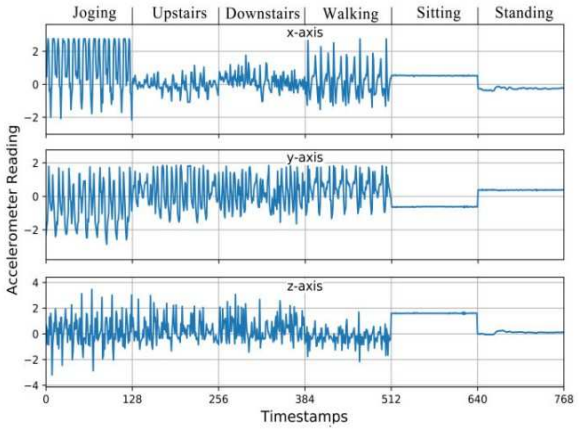


Fig. 2. Activities represented in the form of signals [36]

Although the above models have presented good quality results by understanding and recognizing the actions, but, the overall structure of the network is very complex. Moreover, these techniques have a large number of parameters that can lead to high cost and low speed, especially, when used for practical real-time applications. Therefore, simple LSTM network architecture is proposed in this work that provides high accuracy even with a small number of parameters. This approach reduces the complexity, training time and increases the speed. The resulting parameters are quite promising in comparison to the available research work.

III. Dataset

The dataset used in this research work is Wireless Sensor Data Mining (WISDM) [36]. This is the public dataset and taken from its official website. This data is collected with the help of an accelerometer sensor which is embedded inside the smartphone. It has location details of three-dimensional values x, y and z. It consists of six activities (attributes) that are performed by the individuals at different positions and for separate periods. These activities are coming Downstairs(down), Sitting(sit), Jogging (jog), going Upstairs(up), Standing(std), and Walking(walk). A pie chart shows the percentage division of total samples affiliated with each activity. If we look at the pie chart in Fig. 1. we can easily see that the distribution of total samples for each activity is irregular. The total number of samples in this dataset is 1,098,207. Walking activity has 39% of total data: highest

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:33:43 UTC from IEEE Xplore. Restrictions apply.

among all the activities and standing activity has 4% of total data: representing lowest contribution. So, we can say that the WISDM dataset is unbalanced.

Fig. 2. is representing the visualization of activities in the form of waveforms with 128 points for each activity [36]. Each activity has 3 signals for 3 dimensions (x, y and z). Each signal is plotted with a sampling rate of 50 Hz and represented as a time-series data for each dimension. To remove the noise from the signals we have used the Butterworth filter with a frequency of 20 Hz to get precise results.

IV. Methodology

To improve the accuracy and reduce the complexity we have proposed two models in our research. In the first model, we present the CNN model and the applied number of layers resulting in high accuracy. In the second model, we discuss the LSTM layer and its effect on accuracy.

A. Convolutional Neural Network(CNN)

CNN got attention in recent times due to its ability to extract features from the input and optimize itself while learning. CNN is an impressive and promising form of Artificial Neural Network (ANN). CNN involves the convolution layer, pooling operation and output layer. Kernels of different sizes are used for extracting the features. Kernels are in the form of a matrix that depicts some unique feature that is to be extracted from the input after convolving it with the input and provides output as a feature map.

We have used four layers of convolution with different kernel sizes. The size of the kernel defines the depth of extraction of features from the input. ReLU function is used as a non-linear activation function during the convolution operation. Using large-size kernels results in the deeper extraction of features which leads to overfitting. Overfitting increases the generalization error which results in poor performance of the model on new data. To reduce this snag, we have used dropout layers in between the convolution layers. Dropout gives the regularization effect to our model. Dropout means 'dropping out' the connection of neurons from learning at the time of training. The output layer has Fully Connected Layer (FCL) and a Soft-Max layer. FCL has 32 neurons in the first layer and 64 neurons in the second layer. Soft-max layer has 6 neurons representing the six activities. Soft-max is a classifier that maps the value of output in terms of probability vector that represents the probability of classes. The mathematical expression of the ReLU function and soft-max function is shown below [37]:

$$f(x) = \max(0, x) \tag{1}$$

$$Q_i = \frac{e^{a_i}}{\sum_{p=1}^{N} e^{a_p}} \tag{2}$$

B. Long Short-Term Memory(LSTM)

RNN is another unique form of ANN which can handle the 'temporal-behaviour' of the input. RNN algorithm is a very strong and promising algorithm because it has internal memory to store the information [38]. The unfolded architecture of RNN is shown in Fig. 4. The memory blocks in the architecture can store prior states and with the help of these states they can alter the present state. RNN provides best

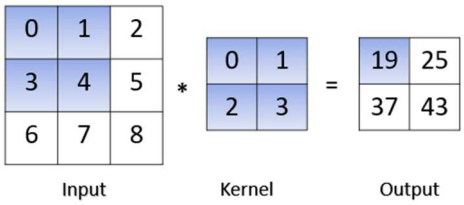


Fig. 3. Convolution

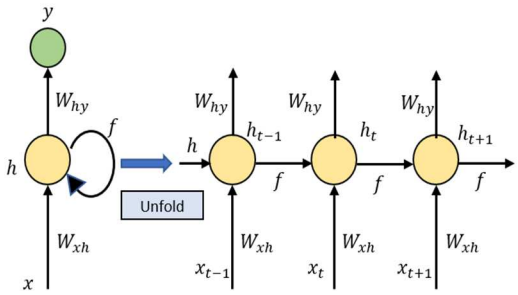


Fig. 4. Figure showing Unfolded version of RNN

results when applied on sequential data. Mathematical expression of present hidden state is shown as [37]:

$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t) \tag{3}$$

here t represents the present time state, $t-1$ is the previous time state and $t+1$ is the future time state. The hidden state at $t-1$ i.e., h_{t-1} is stored in the memory and acts as input at time t. At time t, the inputs are: x_t (current input) and h_{t-1} (previous hidden state). Outputs at time t are: y_t (current output) and h_t (current hidden state). Hyper-tangent function (tanh) is used as the non-linear activation function. W_{hy} *and* W_{xh} are the weights associated with h_t and x_t.

RNN has a major issue of vanishing gradient due to which it cannot handle the long-term dependency of data. LSTM is one of the variants of RNN which can handle the issue of vanishing gradient. LSTM has memory-blocks and these blocks further have the memory cells. In addition to this, different gates are present inside this network. It contains three types of gates in the form of the multiplication-units-input gate, output gate and forget gate [34].

- **Forget gate**: it removes the irrelevant data and decides which information is to be passed to the cell state.
- **Input gate**: this gate decides which new information is to be passed through the cell state inside the network.
- **Output gate**: this gate gives filtered output using the hyper tangent function.

The activation function applied at the gate is 'logistic-sigmoid'. Activation functions at the input of cells are 'hyper-tangent' or 'logistic-sigmoid'. The memory-architecture is shown in Fig. 5. It represents the flow of information within

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:33:43 UTC from IEEE Xplore. Restrictions apply.

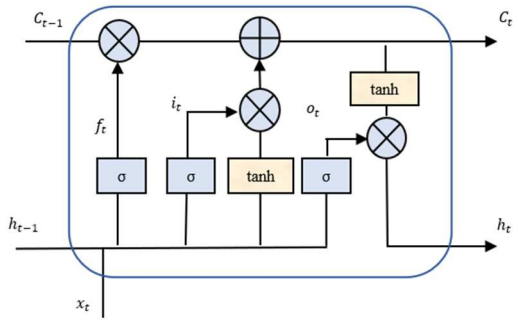


Fig. 5. LSTM internal flow

the network. The mathematical expressions for the activation functions are shown below [37]:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (4)$$

$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + b_f\right) \quad (5)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (6)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (7)$$

$$h_t = o_t \odot c_{t-1} + \tanh(c_t) \quad (8)$$

where σ is the logistic sigmoid function, $\odot$ represent element-wise-multiplication and here, i_t, o_t, f_t, and c_t are input gate, forget gate, memory cell and output gate vector functions (activation functions) respectively. b_i, b_o, b_f and b_c are the bias terms and W_{ab} is representing the weight matrix from a to b.

In this research, two LSTM layers are applied with 32 neurons (memory-cells). Along with this, batch normalization and dropout layers are added. The batch normalization layer is used to get stabilized output by normalizing the data during the training process. Two LSTM layers are giving good results i.e., very high accuracy. By increasing the number of layers, accuracy does not change much but the complexity of the model increases due to the increased number of parameters.

V. Experimental Results

In our experiment, we have used WISDM – public dataset. This dataset is a time-series dataset and we have considered only a part of it. To segment the data from continuous data, a 128-length sliding-window is used with a step size of 64. Training data and test data are split in a ratio of 7:3. Keras with TensorFlow 2.0 is used as an API in python. Initialization of weights and biases are done with the values selected randomly by TensorFlow. The loss function is defined as the difference between the true value and the predicted value. The cross-entropy loss function is used because it yields the faster convergence of the model. Adam optimizer is used to optimize the loss function.

TABLE I. RESULTS FROM THE STUDY

No. of layers	Total Parameters
One LSTM layer	13,622
Two LSTM layers	21,922

Adam optimizer is the "Adaptive Moment Estimation" and is a combination of Adagrad and RMSprop. Adam-optimizer is computationally efficient and requires less memory. The CNN model is evaluated for 100 epochs and LSTM model is evaluated for 200 epochs. An experiment is done on different batch sizes, different dropout values and at different epochs to check the accuracy and loss for the models. The epoch values are selected such that our model converges below that epoch value, and after that, there is no use of increasing the number of epochs. Our model provides good accuracy and F1 score without suffering from overfitting at a batch size of 64, dropout=0.5, 100 epochs (for CNN model) and 200 epochs (for LSTM model). This is the optimal batch size as when we increase it, it leads to overfitting of the model and the convergence is slow. Moreover, when we reduce the batch size it leads to underfitting, is very unstable and behaves like a noisy model. A small batch size model will converge very fast but is not able to locate the global minima. For chosen parameters, the CNN model obtained an accuracy of 95% and the LSTM model obtained 97.5% accuracy.

Fig. 6 represents the confusion matrix of the CNN model and we can see that model got confused between "walking" and "standing". The accuracy and loss curves are shown for training and validation data in Fig. 7-8 and Fig.10-11. Fig. 7 and Fig. 8 represent the accuracy and loss for the CNN model and it converges at 30 epochs. Validation loss is slightly high in this case but is acceptable as the model is converging smoothly. Fig. 9 represents the confusion matrix, Fig. 10 and Fig. 11 show the accuracy and loss for the LSTM model. The confusion matrix of the LSTM model shows that model got confused between "going downstairs" and "coming upstairs". The LSTM model is converging after 75 epochs. These results are quite promising in comparison to other deep learning models applied on the WISDM dataset.

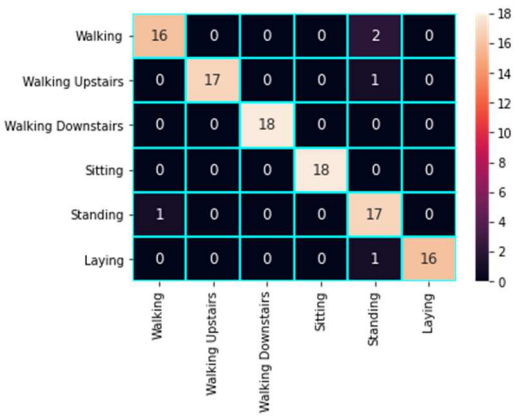


Fig. 6. Classification Confusion Matrix for CNN model

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:33:43 UTC from IEEE Xplore. Restrictions apply.

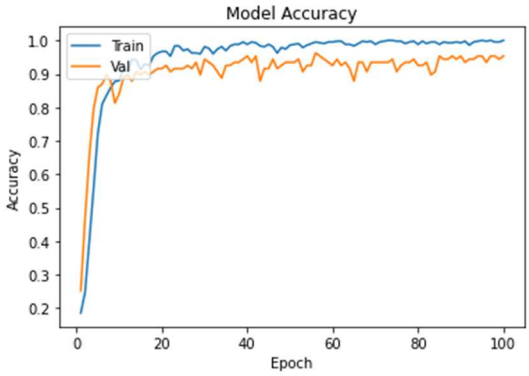


Fig. 7. CNN model accuracy

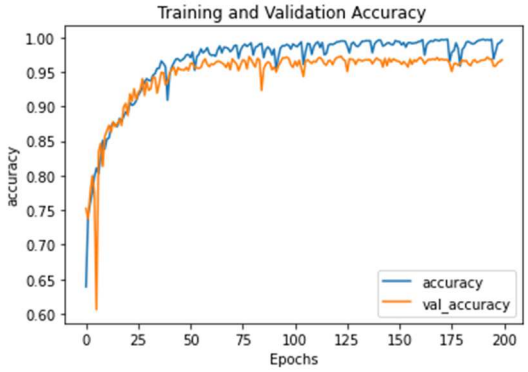


Fig. 10. LSTM model Accuracy

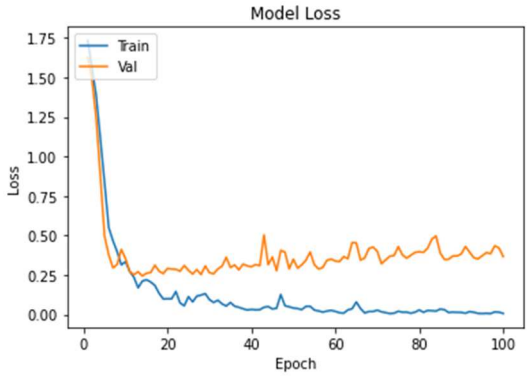


Fig. 8. CNN model Loss

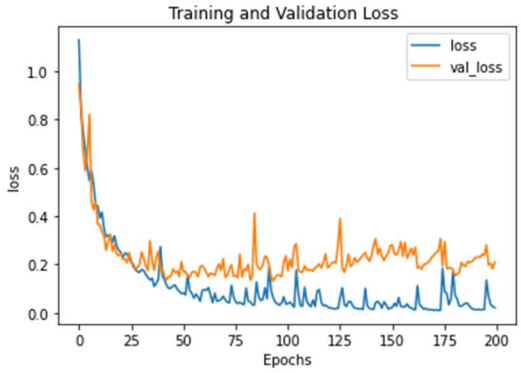


Fig. 11. LSTM model Loss

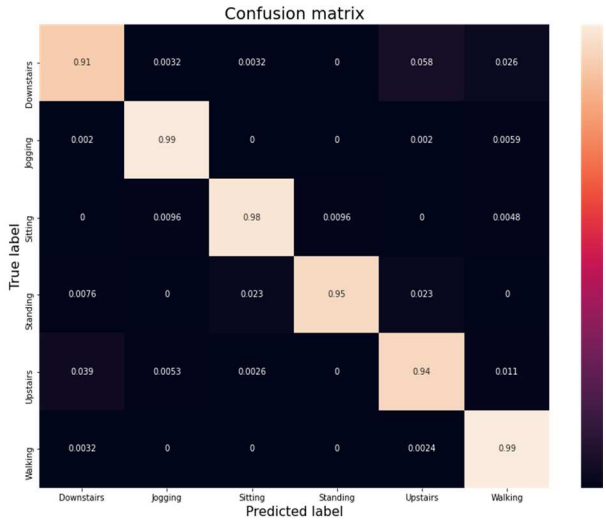


Fig. 9. Classification Confusion Matrix for LSTM model

VI. Conclusion

A smartphone based intelligent human activity recognition system is proposed in this work. The system collected time-series signals using a sensor embedded into a smartphone. The data was preprocessed and converted into a suitable form to apply the deep learning techniques. The model was trained and validated with a 2-D CNN classifier and LSTM network and they classify six different human activities.

The proposed system has high accuracy and low loss. The complexity of the proposed LSTM architecture is low because only two layers are used. Moreover, changes are made to the parameters like dropout, batch size and number of epochs to achieve good results. Accuracy up to 97.5% is achieved using the proposed LSTM network. The proposed LSTM model is found to be more accurate in comparison to other state-of-the-art models available in the open literature. In future work, we will consider more attributes to develop a real-time HAR system. We will also carry out a study to combine different deep learning techniques to further improve the model.

References

[1] Cho Nilar Phyo,Thi Thi Zin , Pyke Tin, “Deep Learning for Recognizing Human Activities Using Motions of Skeletal Joints”, IEEE Transactions on Consumer Electronics, Volume: 65, Issue: 2, May 2019.

[2] J. Wang, Y. Chen, S. Hao, X. Peng and L. Hu, "Deep learning for sensor-based activity recognition: A survey", *Pattern Recognit. Lett.*, vol. 119, pp. 3-11, Mar. 2019.

[3] P. Vepakomma, D. De, S. K. Das, and S. Bhansali, ‘‘A-wristocracy: Deep learning on wrist-worn sensing for recognition of user complex activities,’’ in Proc. IEEE 12th Int. Conf. Wearable Implant. Body Sensor Netw. (BSN), Jun. 2015, pp. 1–6.

[4] Y. Kim and B. Toomajian, ‘‘Hand gesture recognition using micro-Doppler signatures with convolutional neural network,’’ IEEE Access, vol. 4, pp. 7125–7130, 2016

[5] N. Y. Hammerla, S. Halloran, and T. Ploetz, ‘‘Deep, convolutional, and recurrent models for human activity recognition using wearables,’’ 2016, arXiv:1604.08880.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:33:43 UTC from IEEE Xplore. Restrictions apply.

[6] J. Qin, L. Liu, Z. Zhang, Y. Wang, and L. Shao, ''Compressive sequential learning for action similarity labeling,'' IEEE Trans. Image Process., vol. 25, no. 2, pp. 756–769, Feb. 2016.

[7] Lara, O.D., Labrador, M.A., 2013. "A survey on human activity recognition using wearable sensors", IEEE Communications Surveys Tutorials 15 (3), 1192e1209. 10.1109/SURV.2012.110112.00192.

[8] Pfeifer, M., Voelker, B., 2015. "Sensors in human activity recognition". In: Seminar Course.

[9] Chen, C., Jafari, R., Kehtarnavaz, N., 2017. "A survey of depth and inertial sensor fusion for human action recognition". Multimed. Tools Appl. 76 (3), 4405e4425.

[10] Ameur, S., Ben Khalifa, A., Bouhlel, M.S., 2016. "A comprehensive leap motion database for hand gesture recognition". In: 2016 7th International Conference on Sciences of Electronics. Technologies of Information and Telecommunications (SETIT), pp. 514e519. 10.1109/SETIT.2016.7939924

[11] M. Cornacchia, K. Ozcan, Y. Zheng, and S. Velipasalar, ''A survey on activity detection and classification using wearable sensors,'' IEEE Sensors J., vol. 17, no. 2, pp. 386–403, Jan. 2017.

[12] K. Yatani and K. N. Truong, ''BodyScope: A wearable acoustic sensor for activity recognition,'' in Proc. ACM Conf. Ubiquitous Comput. (UbiComp), 2012, pp. 341–350.

[13] B. Cagliyan, C. Karabacak, and S. Z. Gurbuz, ''Human activity recognition using a low cost, COTS radar network,'' in Proc. IEEE Radar Conf., May 2014, pp. 1223–1228.

[14] X. Yang and Y. Tian, ''Super normal vector for human activity recognition with depth cameras,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 5, pp. 1028–1039, May 2017.

[15] H. Wang and C. Schmid, ''Action recognition with improved trajectories,'' in Proc. IEEE Int. Conf. Comput. Vis., Dec. 2013, pp. 3551–3558.

[16] H. Wang, A. Klaser, C. Schmid, and C.-L. Liu, ''Action recognition by dense trajectories,'' in Proc. CVPR, Jun. 2011, pp. 3169–3176.

[17] L. Chen, C. D. Nugent, and H. Wang, ''A knowledge-driven approach to activity recognition in smart homes,'' IEEE Trans. Knowl. Data Eng., vol. 24, no. 6, pp. 961–974, Jun. 2012.

[18] A. Jain and V. Kanhangad, ''Human activity classification in smartphones using accelerometer and gyroscope sensors,'' IEEE Sensors J., vol. 18, no. 3, pp. 1169–1177, Feb. 2018.

[19] Sung, J., Ponce, C., Selman, B., Saxena, A., 2011. "Human Activity Detection from Rgbd images". Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence, pp. 47e55.

[20] P. Gupta and T. Dallas, ''Feature selection and activity recognition system using a single triaxial accelerometer,'' IEEE Trans. Biomed. Eng., vol. 61, no. 6, pp. 1780–1786, Jun. 2014.

[21] E. Fullerton, B. Heller, and M. Munoz-Organero, ''Recognizing human activity in free-living using multiple body-worn accelerometers,'' IEEE Sensors J., vol. 17, no. 16, pp. 5290–5297, Aug. 2017.

[22] L. Bao and S.S. Intille, "Activity Recognition from User-Annotated Acceleration Data," In: Ferscha A., Mattern F. (eds) Pervasive Comput., LNCS, vol. 3001, 2004.

[23] J. Mantyjarvi, J. Himberg and T. Seppanen, "Recognizing human motion with multiple acceleration sensors," in Proc. IEEE International Conference on Systems, Man and Cybernetics, vol. 2, pp. 747-752, 2001.

[24] E. M. Tapia et al., "Real-Time Recognition of Physical Activities and Their Intensities Using Wireless Accelerometers and a Heart Rate Monitor," in Proc. 11th IEEE Int. Symp. Wearable Comput., 2007, pp. 37-40.

[25] A. Jain and V. Kanhangad, ''Human activity classification in smartphones using accelerometer and gyroscope sensors,'' IEEE Sensors J., vol. 18, no. 3, pp. 1169–1177, Feb. 2018.

[26] N. Jalloul, F. Poree, G. Viardot, P. L'Hostis, and G. Carrault, ''Activity recognition using complex network analysis,'' IEEE J. Biomed. Health Informat., vol. 22, no. 4, pp. 989–1000, Jul. 2018.

[27] J. Margarito, R. Helaoui, and A. M. Bianchi, ''User-independent recognition of sports activities from a single wrist-worn accelerometer: A template-matching-based approach,'' IEEE Trans. Biomed. Eng., vol. 63, no. 4, pp. 788–796, Apr. 2016.

[28] C. Zhu and W. Sheng, ''Wearable sensor-based hand gesture and daily activity recognition for robot-assisted living,'' IEEE Trans. Syst., Man, Cybern. A, Syst. Humans, vol. 41, no. 3, pp. 569–573, May 2011.

[29] G. Bieber, J. Voskamp, B. Urban, "Activity Recognition for Everyday Life on Mobile Phones," In: Stephanidis C. (eds) UAHCI. Intelligent and Ubiquitous Interaction Environments, LNCS, 2009, vol. 5615.

[30] Carós JS, Chételat O, Celka P, Dasen S, CmÃral J. "Very low complexity algorithm for ambulatory activity classification". European Medical & Biological Engineering Conference and IFMBE European Conference on Biomedical Engineering. 2005.

[31] Maurer, U., Smailagic, A., Siewiorek, D.P. and Deisher, M., 2006, "Activity recognition and monitoring using multiple sensors on different body positions", In Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on (pp. 4-pp). IEEE.

[32] Wang S, Yang J, Chen N, Chen X, Zhang Q., "Human activity recognition with user-free accelerometers in the sensor networks", International Workshop on Wearable and Implantable Body Sensor Networks (BSN'06). 2005;2:1212-1217.

[33] Y. Zheng, Q. Liu, and E. Chen, ''Time series classification using multi-channels deep convolutional neural networks,'' in Proc. Int. Conf. Web-Age Inf. Manage. Cham, Switzerland: Springer, 2014, pp. 298–310.

[34] F. Ordóñez and D. Roggen, ''Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition,'' Sensors, vol. 16, no. 1, p. 115, 2016.

[35] Y. Lin, J. Le Kernec, S. Yang, F. Fioranelli, O. Romain, and Z. Zhao, ''Human activity classification with radar: Optimization and noise robustness with iterative convolutional neural networks followed with random forests,'' IEEE Sensors J., vol. 18, no. 23, pp. 9669–9681, Dec. 2018.

[36] Kun xia , jianguang huang , and hanyu wang, "LSTM-CNN Architecture for Human Activity Recognition",IEEE Access, vol.8,2020.

[37] Sheng yu , Li xie , Lin liu , and Daoxun xia ," Learning Long-Term Temporal Features With Deep Neural Networks for Human Action Recognition", IEEE Access 10.1109/ACCESS. 2019.2962284.

[38] A Sherstinsky ,"Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network", Physica D 404 (2020) 132306

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:33:43 UTC from IEEE Xplore. Restrictions apply.

2021 2nd International Conference for Emerging Technology (INCET)
Belgaum, India. May 21-23, 2021

CNN-based Person Recognition System for Masked Faces in a post-pandemic world

Parangat Mittal
Dept. of Electronics & Communication Engineering
Delhi Technological University
Delhi, India
parangatmittal1@gmail.com

Kartikeya Pandey
Dept. of Electronics & Communication Engineering
Delhi Technological University
Delhi, India
kartiikey1711@gmail.com

Peeyush Tawani
Dept. of Electronics & Communication Engineering
Delhi Technological University
Delhi, India
tawanipeeyush@gmail.com

Prof. Rajesh Rohilla
Dept. of Electronics & Communication Engineering
Delhi Technological University
Delhi, India
rajesh@dce.ac.in

2021 2nd International Conference for Emerging Technology (INCET) | 978-1-7281-7029-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/INCET51464.2021.9456416

***Abstract*—With everyone now covering their faces with a face mask to avoid contagion of the COVID-19 virus, it becomes an increasingly difficult challenge for the facial recognition systems to identify people wearing masks. Algorithms devised before the pandemic for facial recognition often fail in this context, and consequently a need to understand the working of facial recognition algorithms when presented with occluded faces arises. This project aims to develop a new lightweight Convolutional Neural Network-based algorithm to resolve this issue. The proposed model gives a comparable accuracy with similar models developed in the past. Further, the proposed algorithm is used to create a robust system to ensure adherence of COVID-19 protocols in a real-world environment.**

Keywords— Artificial Intelligence, Convolutional Neural Networks, Face Mask Detection, Masked Face Recognition

I. INTRODUCTION

COVID-19 is an unprecedented crisis leading to a huge number of casualties and security concerns. Several studies on this new virus have proven that the act of wearing a face mask by a healthy as well as an infected person considerably reduces the transmission of this virus. Many guidelines have been issued regarding the mandatory use of masks in public places as well as maintaining a certain distance between people but the main cause of concern is how to enforce these guidelines in a real-world setup where the authorities are responsible for the general health and safety. Another cause of concern is that since the virus is highly communicable, the standard tactile security protocols in any organization such as biometric systems based on fingerprints and passcodes are not safe.

Face recognition is one of the best methods to deal with the above-mentioned problems as it does not require any physical touch and once the person has been identified, he could easily be notified if the mandatory guidelines are not being followed. The problem which arises here is that the face mask makes face recognition a very troublesome task since parts of the face are concealed.

An important difference that needs to be noted at this point is that between Face Mask Detection and Masked Face Recognition. The former indicates whether or not someone is wearing a face mask, while the latter is indicative of who is wearing the face mask. The two must not be confused and must be dealt with separately.

Wearing a face mask causes the following problems:

- Criminals and Fraudsters commit various crimes without being recognised, seeking the perfect cover of a face mask.
- Facial authentication becomes a very tough task when a major portion of one's face is concealed by a mask.
- Present-day face recognition algorithms are not efficient when treated on someone wearing a mask which fails to provide the whole facial image for description.
- Exposing merely the periocular region poses a tough challenge to the task of face recognition.

Due to these problems, face masks have adversely affected the performance of current face recognition methods [1]. While occlusion-resistant Convolutional Neural Networks are suited best for this task, concealed parts of the face make it somewhat difficult for distinct recognition [2]. To increase the reliability of the system, image samples of an individual with and without a face mask can been incorporated in the facial recognition systems to uniquely recognize a person without their knowledge or physical interaction.

In this paper, we are proposing a complete methodology to tackle all the above-mentioned issues. We are proposing a person identification system for detecting and recognising masked faces which would replace the existing identification systems in a post-pandemic world. The proposed system is designed to tackle the following tasks:

1. Detect a person in a frame (Human Detection)
2. Detect if a person is wearing a face mask or not. (Face Mask Detection)
3. Recognise a person who is wearing a mask. (Masked Face Recognition)

The rest of this paper is organized as follows: Section 2 presents the previous/related works. The proposed

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:46:04 UTC from IEEE Xplore. Restrictions apply.

methodology is detailed in Section 3. Experimental results are presented and discussed in the subsequent section (Section 4). Conclusion ends the paper.

II. Previous/Related Works

The main contribution of this work is proposal of new models for Face Mask Detection and Masked Face Recognition, and a complete person identification system incorporating both. We will first review state-of-the-art face detection models, and then move to the datasets available for this new problem.

A. Face Recognition

Face recognition is a computer vision task on which several studies have been performed till now [3]. When Face recognition is compared with other physical recognition systems like biometrics, fingerprint and iris, the former has a much better chance to recognize the identity of a person in a discrete manner. Hence, widely used and has many applications such as security monitoring, forensic analysis, and border control. With the growth of deep neural network-based techniques [4]-[6], and the publicly as well as freely available large-scale face datasets [7],[8], face recognition performance has improved by leaps and bounds [9],[10]. Performance of all the deep learning models as well as State-of-the-Art techniques is barely satisfactory when facing challenges in the form of varying facial expressions, poor resolution, diverse illumination, large-pose differences and occlusion. A summary [11] of various state-of-the-art networks has been presented in Table 1.

TABLE I. State-of-the-art Face Recognition Models

Method	Architecture	Accuracy
FaceNet [12]	GoogleNet-24	99.63 %
DeepFace [13]	AlexNet	97.35 %
DeepID2 [14]	AlexNet	99.15 %
DeepID3 [15]	VGGNet-10	99.53 %
VGGFace [16]	VGGNet-16	98.95 %
SphereFace [17]	ResNet-64	99.42 %

The ongoing pandemic has also affected these face recognition models adversely. All people are forced to wear a Face Mask to curb the spread of the virus. Extensive research is being conducted to incorporate the presence of masks. Performance of all the state-of-the-art face recognition models dealing with people without the mask give great results [18] but fail when people are wearing masks. One such study [19] discusses this problem and proposes a method for masked face recognition with an accuracy of 91.3%.

B. Datasets for Face Recognition

Several datasets have been developed over the years to train and evaluate facial recognition models. Earlier datasets for the same problem were very limited and controlled since the image samples were collected manually. Recent datasets like MegaFace [7], CASIA WebFace [20], IMDB-Faces [21], etc. collect the image samples from the internet as well. These datasets, hence, capture the real-world faces and scenarios with a greater precision and detail. The newer datasets also provide a lot of information about the faces. VGGFace2 [22] provides annotations in the form of depth, cross pose, age, sex and ethnicity. Similarly, WIDER FACE [23] provides various attributes like positions, gender, landmarks, and poses.

For the purpose of this paper, facial datasets containing images of individuals wearing masks as well as without masks are required. Not many such datasets were developed prior to the COVID-19 pandemic since this problem arose as a result of the same. The publicly-available datasets for this problem are MAsked FAces (MAFA) [2] dataset, MaskedFace-Net [24] and Real-world Masked Face Dataset (RMFD) [25] consisting of Masked Face Detection Dataset (MFDD), Real-world Masked Face Recognition Dataset (RMFRD) and Simulated Masked Face Recognition Dataset (SMFRD). All these datasets have been compared in Table 2.

TABLE II. Datasets for Masked Face Recognition

Dataset	Number of Images	Key Features
Masked Faces (MAFA) Dataset [2]	35,806	Images contain masks as well as other occlusions
Masked Faces Detection Dataset [25]	24,711	Contains coordinates of masked faces and if mask is not present
Real-world Masked Face Recognition Dataset (RMFRD) [25]	5,000 masked and 90,000 unmasked	Contains masked as well as unmasked faces
Simulated Masked Face Recognition Dataset (SMFRD) [25]	500,000	Masked counterparts of LFW and WebFace datasets
MaskedFace-Net [24]	137,016	Contains information about correctly and incorrectly worn face masks

III. Methodology

In this paper, we present a person recognition system which will aid in recognizing individuals with whatever part of their face is visible due to the face mask. This system comprises a three-step as shown in Fig. 1, with each stage catering to a specific standalone task.

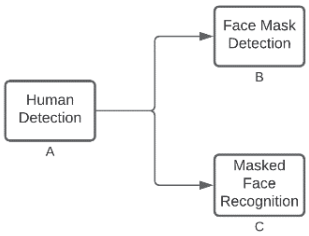


Fig. 1. Stages of the Proposed System.

The first stage, Human Detection, detects the distinct people in a video frame (generally from a CCTV). Once the people are detected, the second stage, Face Mask Detection, checks if the person is wearing a face mask. The last stage, i.e., Masked Face Recognition, confirms the identity of a person who is wearing a face mask for proper authentication.

A. Human Detection

The first function in this task is detecting the people in a video frame, for which already a lot of CNN-based object detection models have been established [26],[27]. Out of all the state-of-the-art methods, YOLO [28] outperforms everyone. This paper uses the same algorithm to achieve the first task.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:46:04 UTC from IEEE Xplore. Restrictions apply.

YOLO v3 model trained on the MS COCO dataset was used in this study, with the detection label "Person". The architecture for the model is shown in Fig. 2.

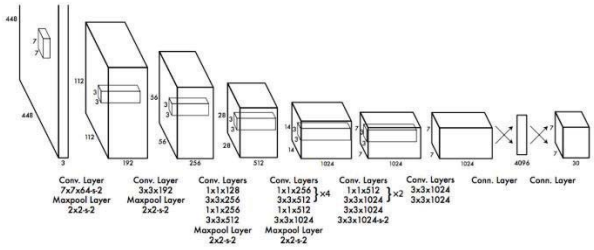


Fig. 2. YOLO v3 Architecture [28].

B. Face Mask Detection

Face mask detection performs the task of verifying whether or not the detected person is covering his face with a mask or not. To achieve this task, we propose the use of a simple Convolutional Neural Network described as follows and illustrated in Fig. 3, which has been trained on images with labels being as masked or unmasked. The output of this stage is binary, since only two classes are being considered – mask or no mask.

The model proposed for this task is as follows:

- **Convolutional Layer 1** consists of 32 kernels of size (3, 3) and a padding of (1, 1), combined with a ReLU operator, and followed by a maxpool with a kernel size of (2, 2).
- **Convolutional Layer 2** consists of 64 kernels of size (3, 3) and a padding of (1, 1), combined with a ReLU operator, and followed by a maxpool with a kernel size of (2, 2).
- **Convolutional Layer 3** consists of 128 kernels of size (3, 3), stride 3 and a padding of (1, 1), combined with a ReLU operator, and followed by a maxpool with a kernel size of (2, 2).
- **Linear Layer** consists of 1024 neurons combined with ReLU.
- **Output Layer** consists of 2 neurons for each of the output class.

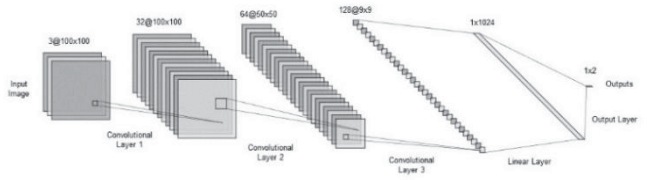


Fig. 3. Proposed Model for Face Mask Detection.

C. Masked Face Recognition

The task of masked face recognition, as explained in the preceding sections, is different from the conventional face recognition task mainly due to unavailability of the complete facial region for processing. Dlib [29] has devised a 68-point facial landmarks model which are considered for CNN-based facial recognition tasks, but in the scope of the current problem maximum 28 facial landmarks can be available in an input image, as shown in Fig. 4. This is with the underlying assumption that the face mask is being properly worn.

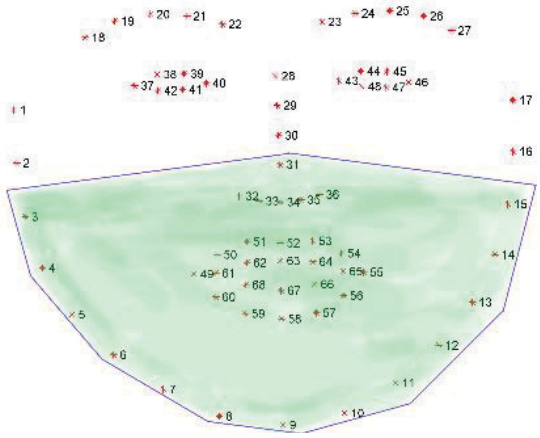


Fig. 4. Facial landmarks in the presence of a mask.

As explained already, a number of occlusion-resistant facial recognition methods have been devised which take into account every kind of obstructions possible. Since we need to account for occlusion only on the lower part of the face, we propose a Convolutional Neural Network as shown in Fig. 5. Since it tackles a specific section obstruction, it is significantly smaller in size and resource-friendly.

The details for the proposed Convolutional Neural Network for this task are as follows:

- **Convolutional Layer 1** consists of 24 kernels of size (10, 10), stride 2 and a padding of (1, 1), combined with a leaky ReLU operator, followed by a maxpool with a kernel size of (2, 2) and a batch normalisation operation.
- **Convolutional Layer 2** consists of 36 kernels of size (9, 9), stride 2 and a padding of (1, 1), combined with a leaky ReLU operator, followed by a maxpool with a kernel size of (2, 2) and a batch normalisation operation.
- **Convolutional Layer 3** consists of 48 kernels of size (8, 8), stride 2 and a padding of (1, 1), combined with a leaky ReLU operator, followed by a maxpool with a kernel size of (2, 2) and a batch normalisation operation.
- **Linear Layer 1** consists of 256 neurons combined with ReLU.
- **Linear Layer 2** consists of 64 neurons combined with ReLU.
- **Output Layer** consists of neurons combined with ReLU equal to the number of output classes.

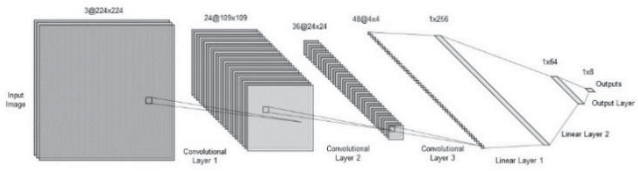


Fig. 5. Proposed Model for Masked Face Recognition

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:46:04 UTC from IEEE Xplore. Restrictions apply.

Due to the limited number of facial landmarks available, the image cannot be processed directly by the network; some level of pre-processing is required in the input image to make it ready for the network.

For a uniform feature extraction from the partially occluded images, it is therefore necessary to align all the images in a similar fashion. Since eyes are the only major facial feature left exposed to the real world while wearing a face mask such that it covers nose, mouth and chin, image alignment can be done using eyes as a reference. In this paper, we consider the line joining the two detected eyes using a Haar Cascade Classifier [30] as the horizontal axis with which the images are aligned.

The pipeline for image processing is as follows (Fig. 6):

1. Detect the face under question and crop it out from the whole image as Region of Interest (ROI)
2. Detect Eyes in the input image and align the image with the line joining both the eyes as horizontal axis
3. Resize and Center Crop the image to 224 x 224 pixels size.

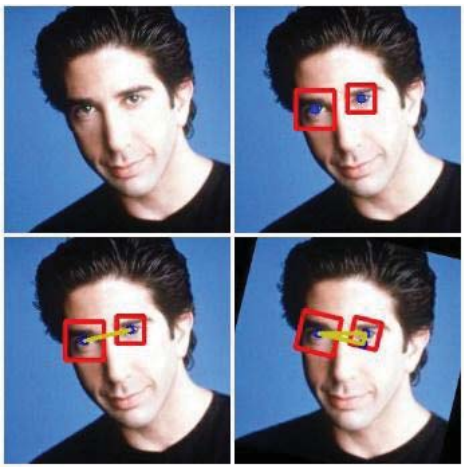


Fig. 6. Recognition Pipeline.

IV. Results And Discussion

A. Dataset

Real-world Masked Faces Dataset (RMFD) [25], a publicly available free dataset, has been used in this study for the tasks of Face Mask Detection and Masked Face Detection. This dataset has laid a virtual mask on all the images of the CASIA-WebFace Dataset [20]. Similar to the CASIA WebFace, our complete dataset also contains 494,414 images over 10,575 identities after face detection.

Images of masked faces from the RMFD [25] and those of unmasked faces from the CASIA-WebFace [20] for each of the identities were included while creation of the training dataset, since in the real-world, not everyone is seen wearing the face mask, detection of which is one of the purposes of this study. The validation or the testing set contained only the images of masked faces. Image samples from the dataset containing masked images and unmasked images of 4 distinct classes have been presented in Fig. 7.



Fig. 7. Unmasked and Masked Image samples.

B. Training and Evaluation

All the images in the dataset were resized to 224 x 224 pixels for uniformity. The grayscale images were removed from all the classes since the models were designed for 3-channel inputs. In this study, the Stochastic Gradient Descent (SGD) and Adaptive Moment Estimation (ADAM) were used as optimizers. Cross Entropy loss was minimized while training the network. The hyper-parameters such as Batch Size, Training Epochs, Learning Rate, Learning Rate Decay Step Size were suitably modified to achieve higher accuracies and lower losses.

The Human Detection Task [28] has already been tested and evaluated since it is the standard detection model. The results for the tasks of Face Mask Detection and Masked Face Recognition are presented in Table 3. Both the proposed models have been compared against a standard state-of-the-art Facenet [12] Inception Resnet v1 model in Table 4.

TABLE III. Results of the Proposed Models

Task	Proposed Model Accuracy	Proposed Model Estimated Size
Face Mask Detection	99.4%	8.23 MB
Masked Face Recognition	97.7%	4.46 MB

TABLE IV. Comparison with Standard Facenet Model

Model	Estimated Size	Floating Point Operations	No. of Parameters
FaceNet Inception Resnet v1	281.39 MB	3,170,217,856	23,486,728
Proposed Model for Masked Face Recognition	4.46 MB	128,352,088	254,476
Proposed Model for Face Mask Detection	8.23 MB	64,460,096	2,193,474

C. Analysis

a) Masked Face Recognition:

In this paper, the proposed CNN for the masked face detection achieves an accuracy of 97.7% which is far better than the earlier state-of-the-art models that were used in the pre-pandemic era. The model has a size of merely 4.46 MB.

It was noted that when the training set consisted of only the masked images of each individual, the model did not achieve a high accuracy in recognition. But when the unmasked images, i.e., images containing the full face, as well as masked images of the individuals were used in the training set, the model achieved stellar accuracy. This allows the feature extractor in the proposed model to

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:46:04 UTC from IEEE Xplore. Restrictions apply.

understand the face completely, and recognise it even when some of it is covered with a mask, as depicted in Fig. 8.

A study [1] in reference to these pre-pandemic state-of-the-art models showed their accuracy to decrease significantly under the current situation. The main problem arising with the earlier state-of-the-art models was that the part of the face occluded by the face mask was being considered by the model for feature extraction. Since no facial feature was available, the model performed poorly. The model proposed in this paper uses only the visible part of the face for extracting the information, resulting in a better performance for this task.

The other models aimed for this task tried to reconstruct the part of the face that was concealed by the mask. Since this is a very resource intensive task such models were unfit and inefficient for real time scenarios. As mentioned above, the proposed model is smaller in size and such a small size makes it suitable for real time applications.

b) Face Mask Detection:

The model proposed for the task of face mask detection in this paper achieves a very high accuracy and is extremely compact in size as well. The accuracy achieved by the face mask recognition is 99.4% with the model size being around 8MB.

The proposed model can be used efficiently in real-life situations due to its compact size. The model detects the face mask only when the mask is worn properly by an individual, i.e., when it completely covers the nose, mouth as well as the chin, as shown clearly in Fig. 9. Whenever any of these facial regions is visible, the model detects it as a "No-Mask" condition, thus making it ideal for ensuring COVID-19 protocols in the real world.

This model acts as a classifier differentiating between the "Mask" and "No-Mask" conditions. It has been trained on the masked images from the RMFRD [26] dataset as a whole nested under the "Mask" condition, while the unmasked images from the CASIA-WebFace [20] dataset as the "No-Mask" condition. Since the images of all the individuals were under a single class, the image samples significantly increased, thus leading to a high accuracy.



Fig. 8. Experimental Results for Masked Face Recognition Task.



Fig. 9. Experimental Results for Face Mask Detection Task.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:46:04 UTC from IEEE Xplore. Restrictions apply.

V. Conclusion

This paper proposes a novel person identification system and details the layout of that using Masked Face Detection and Face Mask Detection. The system can be used in any real-world setup where it is mandatory to ensure the COVID-19 related guidelines. Each component of the system has been tested and verified individually. The above proposed system can also be used at any public place to detect any notorious and suspicious person whose details are available with the authorities and generate an alert for the concerned authorities as well to catch that person. The algorithms for masked face detection and face mask detection proposed in this paper result in a better accuracy than existing algorithms.

Due to the lack of a dataset with all the required parameters, the system could not be implemented as a whole and for future works it is suggested to create a unified dataset which can be used for the above-mentioned task. A dataset of all the individuals with their masked images as well as unmasked images can be created and then the proposed models can be trained on the dataset to achieve the desired results.

References

[1] M. Ngam, P. Grother, and K. Hanaoka, "Ongoing Face Recognition Vendor Test (FRVT) Part 6A: Face Recognition Accuracy with Masks using pre-COVID-19 Algorithms," *NIST IR 8311*, National Institute of Standards and Technology, 2020.

[2] S. Ge, J. Li, Q. Ye and Z. Luo, "Detecting Masked Faces in the Wild with LLE-CNNs," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 426-434, 2017.

[3] L. Best-Rowden and A. K. Jain, "Longitudinal study of automatic face recognition," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(1), pp. 148-162, 2018.

[4] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778, 2016.

[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, pp 1097-1105, 2012.

[6] E. Real, A. Aggarwal, Y. Huang, and Q. V. Le, "Aging evolution for image classifier architecture search," in *AAAI Conference on Artificial Intelligence*, 2019.

[7] I. Kemelmacher-Shlizerman, S. M. Seitz, D. Miller, and E. Brossard, "The megaface benchmark: 1 million faces for recognition at scale," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4873–4882, 2016.

[8] P. J. Phillips, "A cross benchmark assessment of a deep convolutional neural network for face recognition," in *Proceedings of the IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, pp. 705-710, 2017.

[9] Y. Cheng, J. Zhao, Z. Wang, Y. Xu, K. Jayashree, S. Shen, and J. Feng, "Know you at one glance: A compact vector representation for lowshot learning," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pp. 1924–1932, 2017.

[10] L. Tran, X. Yin, and X. Liu, "Disentangled representation learning gan for pose-invariant face recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1415–1424, 2017.

[11] I. Masi, Y. Wu, T. Hassner, and P. Natarajan, "Deep Face Recognition: A Survey" in *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pp 471–478, 2018.

[12] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 815–823, 2015.

[13] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1701–1708, 2014.

[14] Y. Sun, Y. Chen, X. Wang, and X. Tang, "Deep learning face representation by joint identification-verification," in *Advances in Neural Information Processing Systems*, pp. 1988–1996, 2014.

[15] Y. Sun, D. Liang, X. Wang, and X. Tang, "DeepID3: Face Recognition with Very Deep Neural Networks," arXiv:1502.00873, 2015.

[16] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep Face Recognition," in *Proceedings of the British Machine Vision Conference (BMVC)*, pp. 41.1-41.12, 2015.

[17] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, and L. Song, "Sphereface: Deep hypersphere embedding for face recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 212–220, 2017.

[18] M. A. Hmani and D. Petrovska-Delacrétaz, "State-of-the-art face recognition performance using publicly available software and datasets," in *Proceedings of the International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pp. 1-6, 2018.

[19] W. Hariri, "Efficient Masked Face Recognition Method during the COVID-19 Pandemic," Preprint (Version 3) available at Research Square, 2020.

[20] D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Learning Face Representation from Scratch," arXiv:1411.7923, 2014.

[21] F. Wang, L. Chen, C. Li, S. Huang, Y. Chen, C. Qian, and C. Change Loy, "The Devil of Face Recognition is in the Noise," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.

[22] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman, "Vggface2: A dataset for recognising faces across pose and age," in *Proceedings of the IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, pp. 67–74, 2018.

[23] S. Yang, P. Luo, L. C. Change, and X. Tang, "Wider face: A face detection benchmark," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp 5525–5533, 2016.

[24] A. Cabani, K. Hammoudi, H. Benhabiles, and M. Melkemi, "MaskedFace-Net - A Dataset of Correctly/Incorrectly Masked Face Images in the Context of COVID-19", arXiv:2008.08016, 2020.

[25] Z. Wang, G. Wang, B. Huang, Z. Xiong, Q. Hong, H. Wu, P. Yi, K. Jiang, N. Wang, Y. Pei, H. Chen, Y. Miao, Z. Huang, and J. Liang, "Masked face recognition dataset and application," arXiv:2003.09093, 2020.

[26] Z. Zhao, P. Zheng, S. Xu, X. Wu, "Object Detection with Deep Learning: A Review," arXiv:1807.05511, 2019.

[27] L. Jiao, F. Zhang, F. Liu, S. Yang, L. Li, Z. Feng, R. Qu, "A Survey of Deep Learning-Based Object Detection," in *IEEE Access*, vol. 7, pp. 128837–128868, 2019.

[28] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," arXiv:1506.02640, 2016.

[29] Davis E. King, "Dlib-ml: A Machine Learning Toolkit," in *Journal of Machine Learning Research*, vol. 10, pp. 1755-1758, 2009.

[30] P. I. Wilson, and J. Fernandez, "Facial feature detection using Haar classifiers," in *Journal of Computing Sciences in Colleges*, vol. 21, no. 4, pp. 127–133, 2006.

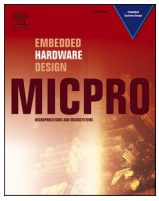
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:46:04 UTC from IEEE Xplore. Restrictions apply.

Microprocessors and Microsystems 85 (2021) 104281

Contents lists available at ScienceDirect

Microprocessors and Microsystems

journal homepage: www.elsevier.com/locate/micpro



Comparative Analysis of the Design Techniques for Low Leakage SRAMs at 32nm

Monica Gupta [a,b], Kirti Gupta [b], Neeta Pandey [a,*]

[a] Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, 110042, India
[b] Department of Electronics and Communication Engineering, Bharati Vidyapeeth's College of Engineering, Delhi, 110063, India

ARTICLE INFO

Keywords:
Leakage reduction
SRAM
Hold mode
Sub-threshold
PVT-variations

ABSTRACT

This paper presents a comprehensive overview of leakage reduction techniques prevailing in Static Random Access Memories (SRAMs) by classifying them in three categories namely latch, bitline and read port. The performance of the techniques are evaluated in terms of leakage reduction capability along with the impact on read performance and hold stability through extensive simulative investigations at 32 nm technology node by taking conventional SRAM cell as reference. Further, as SRAMs are susceptible to inter-die as well as intra-die process variations, the performance at different PVT corners is also captured to demonstrate the efficacy of each technique under PVT variations. It is found that among the techniques used for reducing latch leakages, Multi-threshold CMOS technique possess the highest leakage reduction capabilities followed by Drowsy mode and Substrate-bias techniques. The results also indicate that Negative word line technique is more effective at low supply voltages whereas the Leakage biased bitline technique is more effective at high supply voltages for reducing bitline leakages. Amongst the read port leakage reduction techniques, Stack-effect and Dynamic control of power supply rail techniques are capable of suppressing the leakages at high voltages whereas Virtual cell ground technique is more efficacious at low voltages. The impact of technology scaling on SRAM cell performance with leakage reduction techniques is also studied. For the sake of completeness, suggestions are put forward for adopting a particular technique to address leakages at latch, bitline and read port levels.

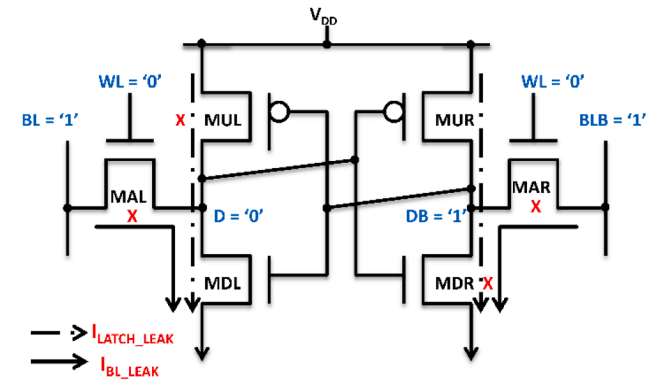
1. Introduction

The power-constrained devices such as mobile phones, tablets etc. require a system with high performance and long battery lifetimes. In order to meet the requirements of such systems, larger on-chip memories are needed. In fact, as per International Technology Roadmap for Semiconductors (ITRS), Static Random-Access Memories (SRAMs) are expected to occupy approximately 94% of the total chip area in near future [1]. This translates into high density designs wherein the large number of SRAM cells are integrated over a smaller chip area [2]. There is a substantial increase in power consumption in such designs due to leakage which is of small magnitude otherwise [2, 3]. Further, the ITRS predicts that the SRAM leakage would have more than 50 % share in total power consumption in the next generation processors [1]. It is pertinent to mention here that sub-threshold leakage current flowing through OFF transistors is one of the major contributors in leakage power [4, 5]. The magnitude of the leakage current varies with the operating mode as set of transistors that remain in OFF state depends on the operation being performed in the SRAM cell.

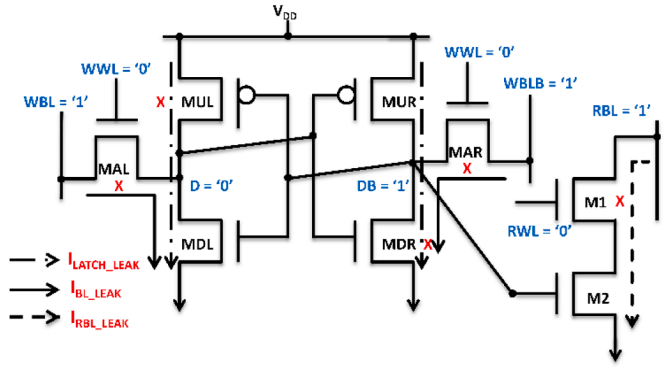
The paper presents a comprehensive overview of the techniques that address leakage reduction in the conventional design of SRAM cell [9-45]. The existing techniques are classified based on their impact within the cell. The first category covers the techniques that aim to lower leakage current that flows from cell supply to ground through non-conducting latch transistors [9-28]. The techniques reducing the bitline leakage current forms the second category [29-36]. As there exists a trade-off between read and write performance of 6T SRAM cell at lower technology nodes [4, 6, 7, 37-45], the techniques adopted in SRAM cells with decoupled read port are also reviewed and are placed in third category. In short, the leakage is addressed at the level of latch, bit line and read port. Further, the performance of a technique is also susceptible to process-supply voltage-temperature (PVT) variations at lower technology nodes [8]. So, the effectiveness of a particular technique depends on its capability to suppress leakage and also its impact on read performance and hold stability under PVT variations. A technique that can provide good leakage reduction capabilities under PVT

* Corresponding author.
E-mail addresses: monica.gupta.me@gmail.com (M. Gupta), Kirtigupta22@gmail.com (K. Gupta), neetapandey@dce.ac.in (N. Pandey).

https://doi.org/10.1016/j.micpro.2021.104281
Received 9 December 2020; Received in revised form 26 March 2021; Accepted 28 May 2021
Available online 10 June 2021
0141-9331/© 2021 Elsevier B.V. All rights reserved.



(a) 6T SRAM cell



(b) 8T SRAM cell [6]

Fig. 1. Leakage currents in conventional SRAM cells

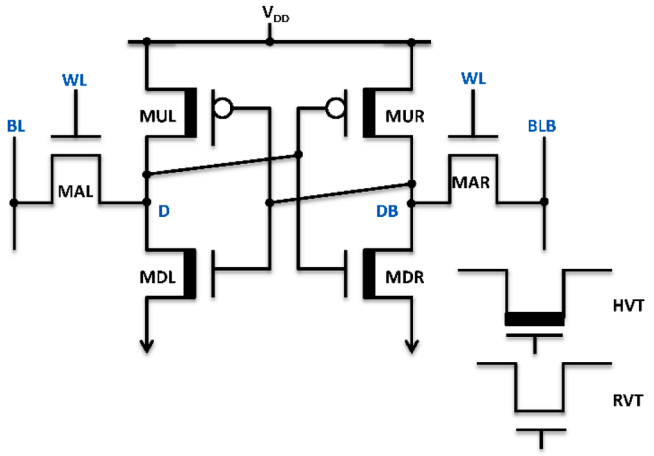


Fig. 2. Schematic of 6T SRAM cell using MTCMOS technique

variations at lower technology nodes is desirable. The leakage current is sensitive to technology scaling, therefore, the efficacy of the technique depends upon the technology node at which the technique is used.

Further, the performance of a technique is also susceptible to process-supply voltage-temperature (PVT) variations at lower technology nodes [8]. So, the effectiveness of a particular technique depends on its capability to suppress leakage and also its impact on other performance parameters under PVT variations. A technique that can provide good leakage reduction capabilities under PVT variations at lower technology nodes is desirable.

The objective of the paper is to quantify the impact of available leakage reduction technique on SRAM cell performance under PVT variation so that a suitable technique may be used at lower technology

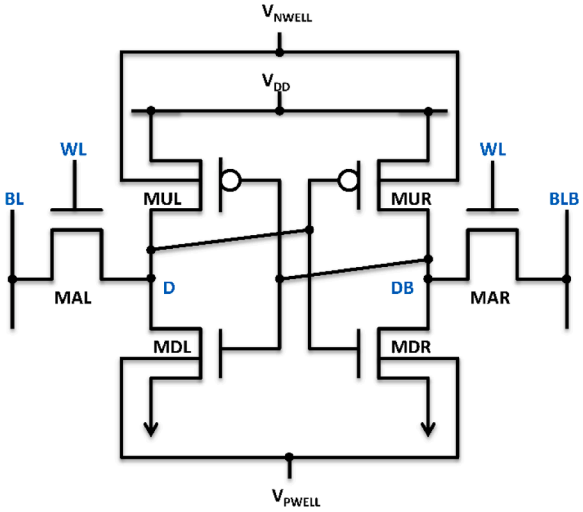


Fig. 3. Schematic of 6T SRAM cell using SB technique

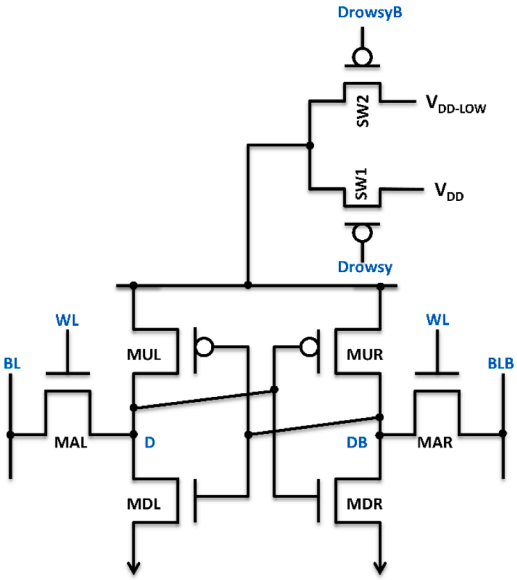


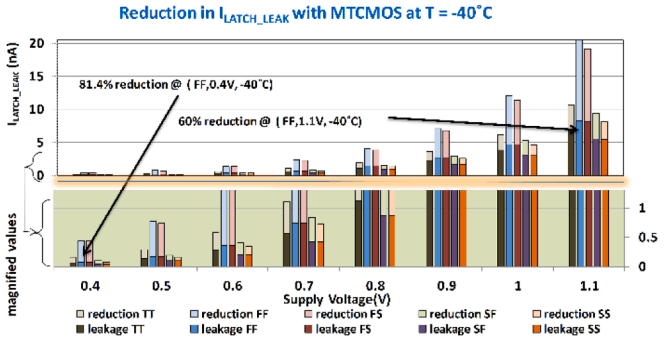
Fig. 4. Schematic of 6T SRAM cell using DM technique

node.

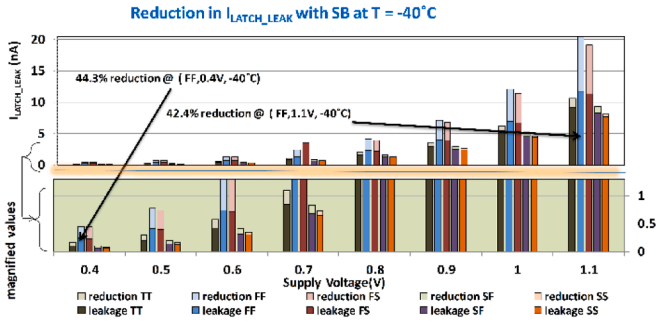
The paper is organized as follows. Section 2 presents the various leakage current components in conventional SRAM cells. Various leakage reduction techniques to address latch leakage, bitline leakage and leakage in read ports are briefly discussed in sub-sections of Section 3. This section also includes findings of extensive simulations carried out to observe the efficacy of a technique in terms of its leakage reduction capability and further impact on read performance and hold stability under PVT variations. A comparative analysis is also presented for techniques falling under each category. Section 4, performs the technology comparison of all the techniques on SRAM cell performance whereas Section 5 finally concludes the paper.

2. Leakage current components in conventional SRAM cells

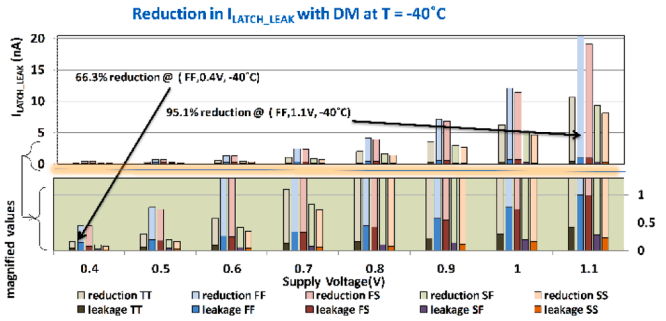
In this section, we briefly describe the various leakage current components existing in conventional structures. A typical SRAM cell stores one bit of data and retains it as long as the power supply is provided. It comprises of a latch and access transistors for storing and accessing the data. The realization of conventional SRAM cell structures



(a) using MTCMOS technique

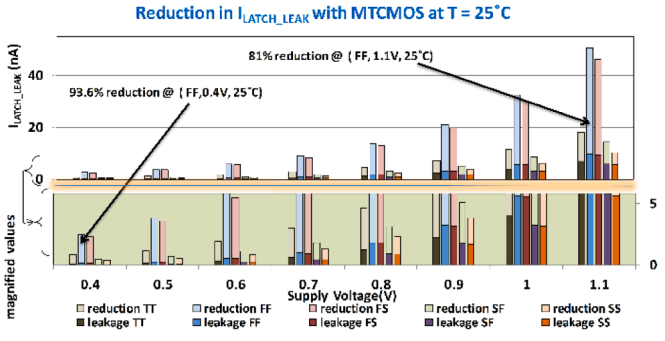


(b) using SB technique

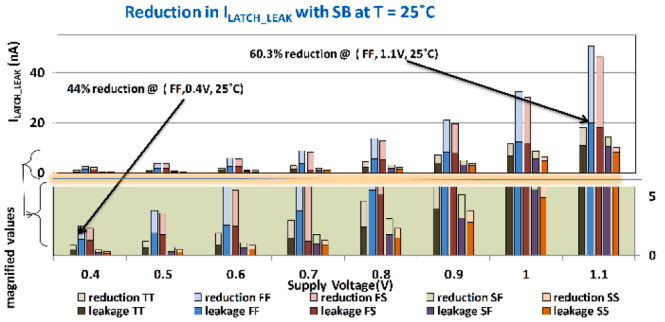


(c) using DM technique

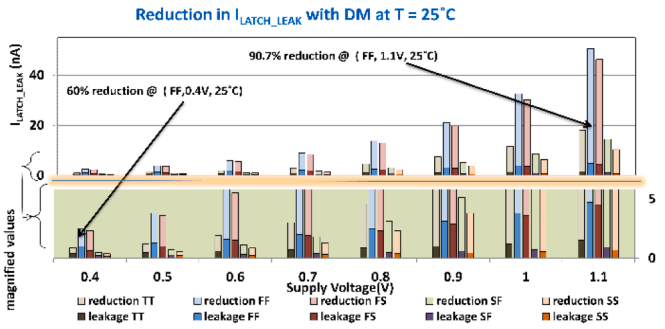
Fig. 5. Effect on Latch leakage current (I_{LATCH_LEAK}) at T = -40°C



(a) using MTCMOS technique



(b) using SB technique



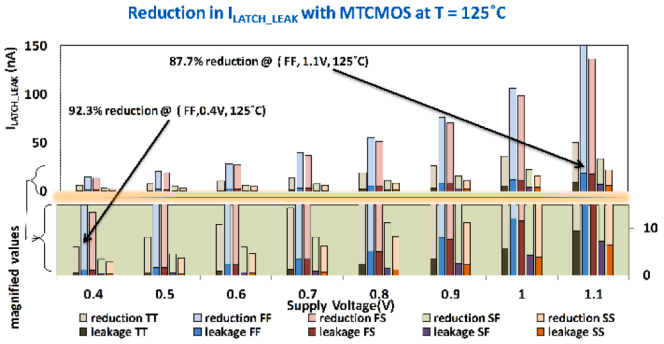
(c) using DM technique

Fig. 6. Effect on Latch leakage current (I_{LATCH_LEAK}) at T = 25°C

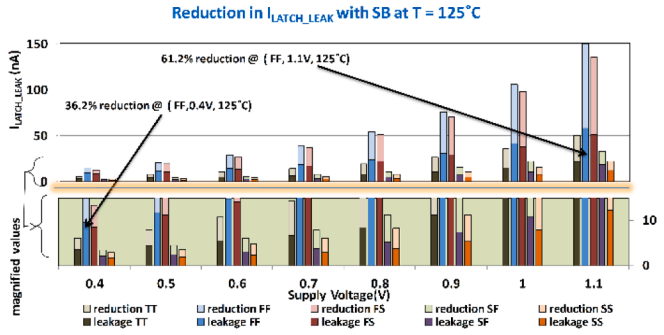
is shown in Fig. 1. The conventional structure (Fig. 1(a)) has six transistors and thus is termed as 6T SRAM cell. The transistors (MUL-MDL and MUR-MDR) forms the latch to store values at two internal nodes, D and DB while the two access transistors (MAL and MAR) enable write/read data operation in the cell. The cell operates in three modes namely write, read and hold to write the new data, to retrieve the data and to preserve it respectively. The complementary bitlines (BL and BLB), wordline (WL) performs the read and write operations via access transistors.

In the conventional design of 6T SRAM cell, there exists a trade-off between read and write performance [4, 6, 7]. Due to the contradictory requirements of strong access transistor and weak latch for faster write operation and opposite for reliable read operation, many alternate SRAM designs are available wherein read port is decoupled. An implementation of 8T SRAM cell [6] with decoupled read port is shown in Fig. 1(b). The transistors (MUL-MDL, MUR-MDR), (MAL-MAR) and (M1, M2) constitute the latch, write and read ports respectively. It exhibits single-ended read in contrast to the differential read in 6T SRAM cell. The control signal, write wordline (WWL), and write bitlines (WBL, WBLB) are used during write while read wordline (RWL) and read bitline (RBL) are used during read operation.

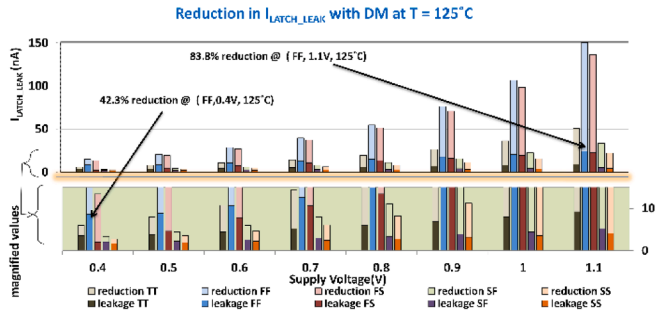
In a SRAM cell, three components of leakage current are identified based on their origin within the cell. The first component is the latch leakage current (I_{LATCH_LEAK}) that flows from cell supply to ground through non-conducting latch transistors. The bitline leakage current (I_{BL_LEAK}) is the second component that flows from BL/BLB to ground through non-conducting access transistors. The third one is defined for cells with decoupled read port and is named as read port leakage current (I_{RBL_LEAK}). A pictorial representation for leakage currents in 6T and 8T [6] SRAM cells storing logic '0' at D during hold mode is shown in Fig. 1 (a) and Fig. 1(b) respectively. In nanometer regime, the leakage current becomes a critical performance parameter as the diffusion of the charge carriers in the source and drain regions is increased which results in a current flow in a non-conducting transistor [11]. This leakage current affects the overall performance of the memory (such as number of SRAM cells connected to sense amplifiers, speed of operation etc.) and makes the cell more sensitive to variations [8]. Therefore, various techniques are used to suppress the leakage current. An investigation on the techniques and their effects on cell performance in nanometer regime are covered in this paper.



(a) using MTCMOS technique



(b) using SB technique



(c) using DM technique

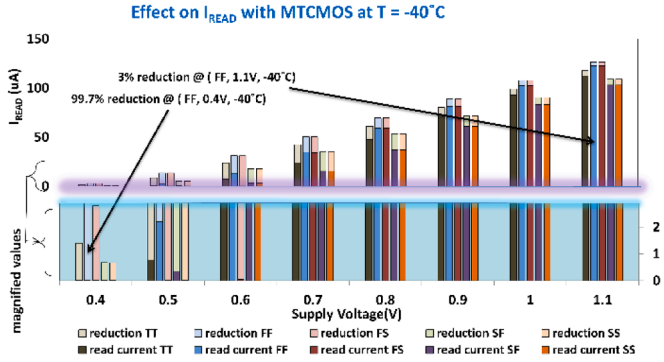
Fig. 7. Effect on Latch leakage current (I_{LATCH_LEAK}) at T = 125°C

3. Leakage reduction techniques for SRAM cells

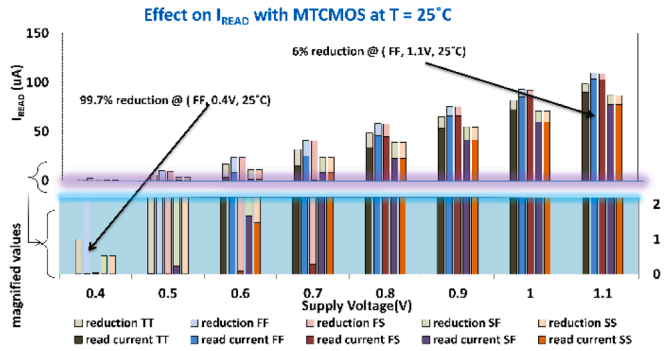
This section overviews the leakage reduction techniques used in SRAM cells. The existing techniques are classified in three categories namely latch leakage reduction techniques, bitline leakage reduction techniques and read bitline leakage reduction techniques on the basis of leakage current origin. The techniques suggest either a circuit level modification in cell topology/multiple supply rails or at device level by using improved MOS transistors. The efficacy of each technique is determined by analyzing its impact on leakage reduction as well as on the read performance and hold stability across different voltage, temperature and process corners. The SRAM cells are simulated using 32 nm CMOS technology parameter with threshold voltage of nMOS ($V_{th,n}$) and pMOS ($V_{th,p}$) as 0.483 V and - 0.452 V respectively.

3.1. Latch leakage reduction techniques [9-28]

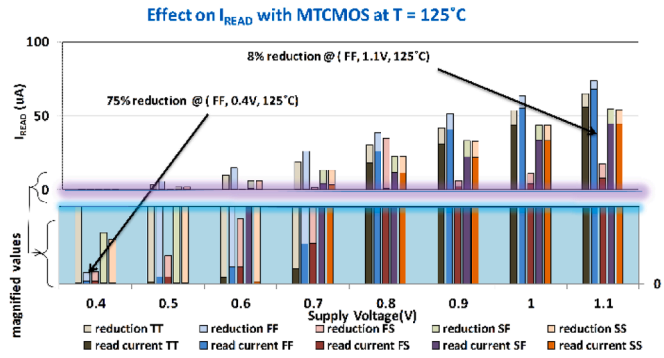
The techniques in this category aim at reducing the leakage current generated within the latch in a SRAM cell. As latch is the integral part of a cell, therefore these techniques can be applied to different SRAM cells to achieve the purpose. A detailed description on the techniques is presented below.



(a) at T = -40 °C



(b) at T = 25 °C



(c) at T = 125 °C

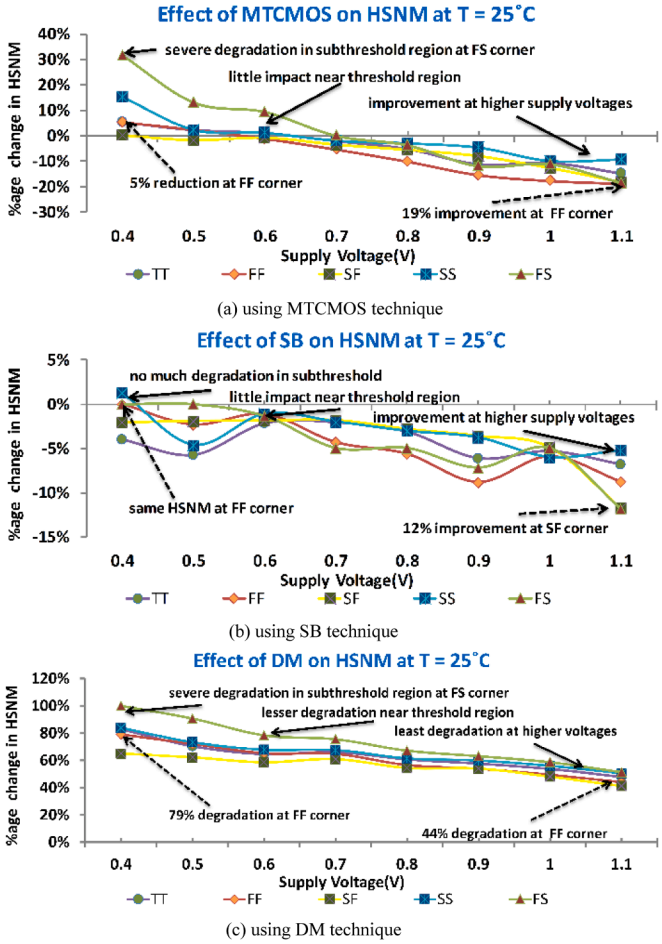
Fig. 8. Effect on Read current (I_{READ}) using MTCMOS technique

3.1.1. Multi-threshold CMOS technique (MTCMOS) [9-12]

This technique suggests the use of high-threshold voltage (HVT) MOS transistors in latch while regular threshold voltage (RVT) transistors to be used in the remaining cell. The presence of HVT MOS transistors lowers the I_{LATCH_LEAK} of the cell. A schematic of 6T SRAM cell using MTCMOS technique is depicted in Fig. 2 where the transistors with bold lines denote HVT MOS transistors. The modification of the threshold voltages is done at fabrication level by changing the doping profiles of the selected transistors yielding a low-leakage SRAM cell.

3.1.2. Substrate-bias technique (SB) [11, 13-16]

This technique presents an alternate way to accomplish the same objective. The threshold voltage of the latch transistor is varied by adjusting the bias voltages of p-wells and n-wells. The n-well voltage (V_{NWELL}) is set higher than V_{DD} while the p-well voltage (V_{PWELL}) is set lower than ground voltage to enhance the body bias effect as shown in Fig. 3. A large offset voltage ($|V_{NWELL} - V_{DD}| = |0\text{-}V_{PWELL}|$) is required to



(a) using MTCMOS technique

(b) using SB technique

(c) using DM technique

Fig. 9. Effect on Hold Static Noise Margin (HSNM) at T = 25°C

cause a significant reduction in I_{LATCH_LEAK}. In hold mode, the non-zero offset voltage increases the threshold voltage while in read or write mode it is set to zero by changing the bias conditions.

3.1.3. Drowsy mode technique (DM) [17-23]

In this technique, the latch transistors in 6T SRAM cell is made to operate in Drowsy Mode wherein the supply voltage is lowered from V_{DD} to $V_{DD\text{-}LOW}$ in order to reduce drain to source voltage across the transistors. This is realized with the help of two MOS-based switches SW1 and SW2 operated using control signals Drowsy and its compliment DrowsyB respectively as shown in Fig. 4. A reduction in I_{LATCH_LEAK} is observed due to lower drain induced barrier lowering effect and reduced gate leakage. It is worth mentioning that the full collapse of supply voltage in 6T SRAM cell using DM technique should be avoided as it would lead to loss of data. The cell supply voltage is generally kept above its data retention voltage (DRV) that is defined as the minimum supply required to maintain its data in a SRAM cell. Additionally, with dynamic supply variation, it is ensured that supply voltage is restored to its optimum value before the contents of the cell are accessed for read or write operation. The different architectures to reduce the area overhead by sharing the floating supply rails amongst the cells on a single column are available in the literature.

3.1.4. Comparative analysis

The SRAM cells using the discussed latch leakage reduction techniques are simulated and their performances are compared with conventional 6T SRAM cell in terms of leakage reduction capability, read performance and hold stability. The performance parameters namely, leakage current, read current and hold static noise margin are measured.

The SRAM cells are simulated at different temperatures (-40 ℃, 25 ℃ and 125 ℃) by varying the supply voltage from 0.4 V to 1.1 V across all process corners (TT, FS, SF, SS, FF; T: Typical, F: Fast, S: Slow) [23] to observe the effect of collective variations. As the leakage is process corner dependent and the ratio of leakage at FF corner to that at SS corner is usually high. So, for fair comparison, the performance of all the techniques is compared at maximum leakage corner (FF). In SB technique the V_{NWELL} and V_{PWELL} of the latch transistors are set at 20 % above V_{DD} and 20 % below ground respectively during hold mode. In the same mode for DM technique, the supply rail is lowered 20 % below

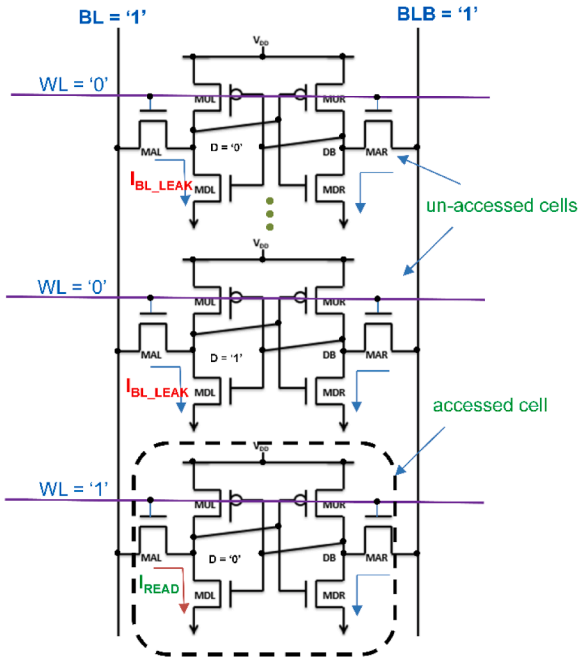


Fig. 10. Flow of I_{READ} in accessed cell and I_{BL_LEAK} in un-accessed cells

Table 1
Performance comparison of latch leakage reduction techniques [15] [24-28]

Techniques	MTCMOS	SB	DM
Leakage	Overall good but more effective at low voltages	Less effective compared to other two and shows better impact at high voltages	Less effective than MTCMOS and shows better impact at high voltages
Read current	Decreases at low voltages	No impact	No impact
Hold stability	Decreases at low voltages	No much impact	Severe degradation at low voltages
PVT tolerant	No	No	No
Suitable temperatures	High temperatures	-40°C to 125°C	-40°C to 125°C
Suitable supply voltages	high voltages	0.4 V to 1.1V	high voltages
Remarks	No area overhead. Requirement of additional implant processing step.	Requirement of twin or triple well processes and extra time to bias the wells properly.	Area overhead. Requirement of multiple supply rails. Extra time to switch between drowsy and non-drowsy mode.

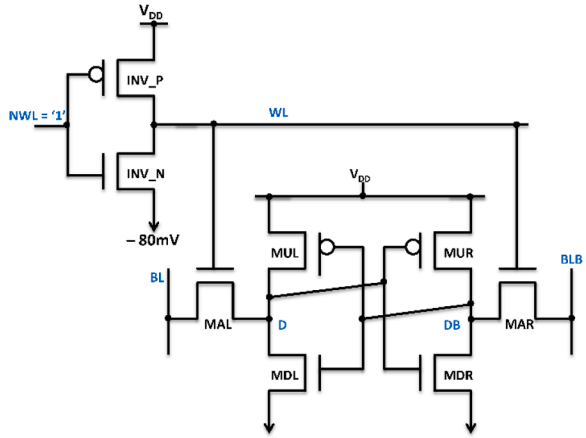


Fig. 11. Schematic of Negative word line (NWL) 6T SRAM cell

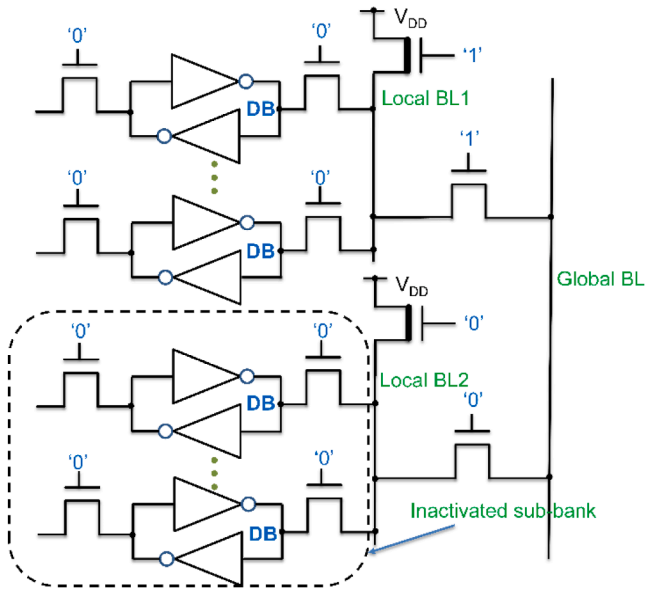


Fig. 12. Inactivated sub-banks with LBB technique in Hierarchical bit line architecture

nominal supply.

3.1.4.1. Leakage reduction capabilities. The reduction in leakage current I_{LATCH_LEAK} obtained for the SRAM cells using different techniques at all operating conditions is plotted in Fig. 5 to Fig. 7. It can be observed in Fig. 5(a) that at -40 ℃ I_{LATCH_LEAK} for SRAM using MTCMOS technique is reduced by 60 % at 1.1 V and 81.4 % at 0.4 V. With SB technique, the percentage reduction in I_{LATCH_LEAK} is 42.4 % at 1.1 V and 44.3 % at 0.4 V showing marginal change (Fig. 5(b)). However, the SRAM cell using DM technique (Fig. 5(c)) shows a high reduction percentage of 95.1 % at 1.1 V which reduces to 66.3 % at 0.4 V. Thus, at high voltages DM technique exhibits the best leakage reduction capabilities while MTCMOS technique is suited for sub-threshold region. At 25 ℃, out of the three techniques only MTCMOS shows improved leakage reduction capabilities with reduction percentages increasing to 93.6 % and 81 % respectively at 0.4 V and 1.1 V as shown in Fig. 6(a). The SB and DM techniques offers reasonable performance at 0.4 V with 44 % and 60 % leakage reduction respectively (Fig. 6(b) & (c)). The leakage reduction using SB technique improves to 60.3 % at 1.1 V while that using DM technique reduces to 90.7 % at the same voltage. Thus, MTCMOS technique shows continuously better performance than others at moderate temperatures particularly in sub-threshold region. As the temperature further increases from 25 ℃ to 125 ℃, the same trend is observed with respect to all the techniques. The MTCMOS technique shows better leakage reduction capabilities over the others at all the voltages as depicted in Fig. 7. Thus, it can be concluded that a SRAM cell designed using MTCMOS technique has improved leakage performance with the increase in temperature especially in sub-threshold region.

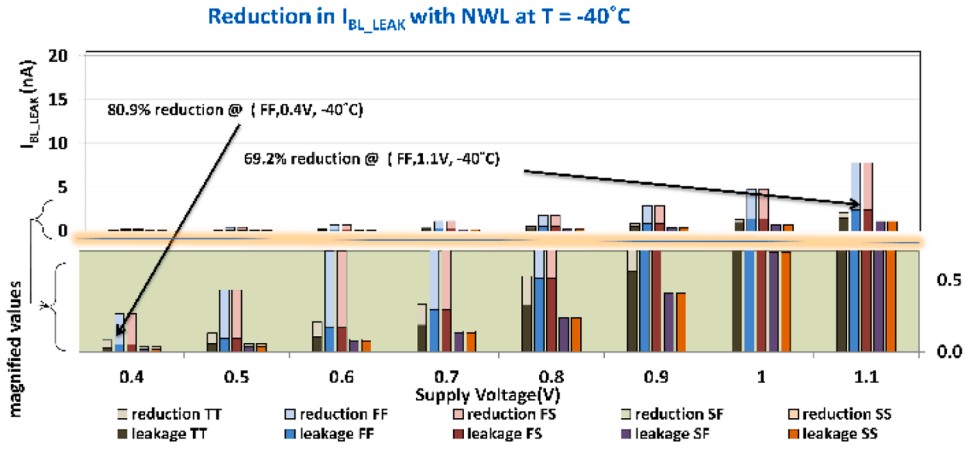
3.1.4.2. Impact on read performance. The read current is measured for the SRAM cells. It may be noted that the SRAM cell with SB technique maintains the same read current values as the conventional cell since the threshold voltage of latch transistors is set to their nominal values during read or write operation by adjusting the Substrate bias voltages. Similarly, DM technique based cell operates at the optimum supply voltage in read/write mode therefore it also shows the same read current values as the conventional cell. The I_{READ} values for MTCMOS based SRAM cell at different temperatures is plotted in Fig. 8. It is observed that I_{READ} is reduced in the sub-threshold region at all temperatures. Alternatively, the impact of technique at higher temperature and voltages shows negligible effect on read current with about 8 % reduction in read current at 1.1 V and FF corner.

3.1.4.3. Impact on hold stability. The HSNM of all the SRAM cells is measured and the percentage change in the HSNM with respect to the conventional cell is plotted in Fig. 9. The MTCMOS based cell (Fig. 9(a)) shows improvement in HSNM values at all corners at 1.1 V with upto 19 % improvement at FF corner. However, as the supply voltage scales down, the HSNM degrades with upto 31 % reduction in the sub-threshold region making the stored data unstable. In SB technique (Fig. 9(b)), significant improvements are observed in HSNM with respect to conventional SRAM cell at all voltages with upto 12 % improvement at SF corner. Even as the supply voltage scales down, no significant degradation of HSNM values is seen at 0.4 V. The dynamic variation of supply voltages in DM technique leads to substantial reduction in HSNM of the cell thereby making cell unsteady in hold mode. The plots in Fig. 9 (c) confirms the preposition. The severe degradation in HSNM of the 6T SRAM cell in the presence of this technique especially in sub-threshold region poses a challenge and makes it unsuitable at lower voltages.

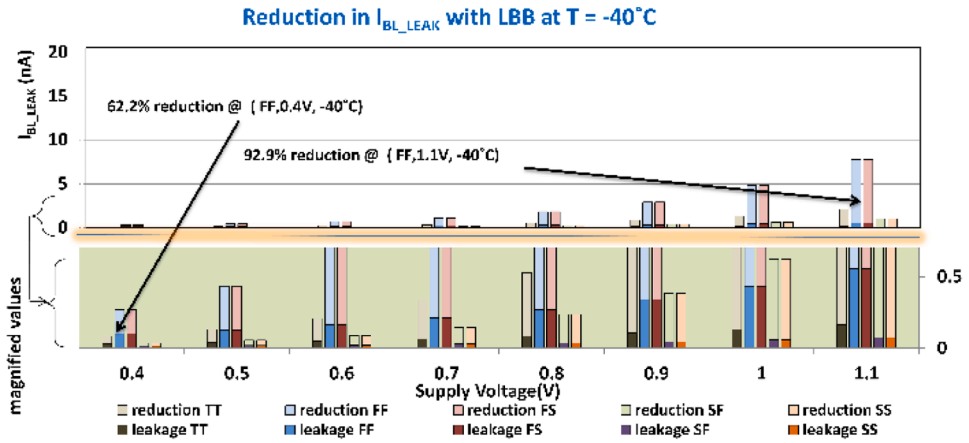
The simulative investigations on latch leakage reduction techniques are summarized in Table 1 for quick access. The techniques have been compared on the basis of read performance and hold stability in addition to the leakage reduction capabilities in order to determine the best suitable operating conditions in which the corresponding technique can be used. It can be observed from the results that MTCMOS and DM techniques are more suitable at high voltages with MTCMOS preferrable at high temperatures. The SB technique, on the other hand, shows moderate performance throughout the supply voltage and temperature range. In addition, none of the technique is tolerant to PVT-variations.

3.2. Bitline leakage reduction techniques [4, 22, 25, 26, 29-35]

A typical SRAM array has memory cells arranged in rows and columns such that all the cells in a particular column perform memory operation through common bitlines (BL, BLB). During a read operation, read current I_{READ} and bitline leakage current I_{BL_LEAK} flow through the accessed cell and the un-accessed cells in a column respectively as depicted in Fig. 10. The value of bitline leakage current I_{BL_LEAK} depends on the data value stored at the corresponding internal nodes (D or DB) in un-accessed cells [4]. It is worth mentioning that the leakage current I_{BL_LEAK} is higher when logic '0' is stored at node D than the case when it stores '1'. As the number of cells sharing the bitline increases, the ratio I_{READ}/I_{BL_LEAK} decreases which in turn leads to higher read failures [22]. This imposes limitation on the number of SRAM cells that may be connected in a particular column in high density SRAMs [26]. The bitline leakage reduction techniques are very useful in such cases.



(a) using NWL technique



(b) using LBB technique

Fig. 13. Effect on Latch leakage current (I_{LATCH_LEAK}) at T = -40°C

In this sub-section, the techniques for reducing the bitline leakage current are covered. The first obvious technique is to downsize the access transistors but it directly impacts the read and write performance of SRAM cell so it is not a preferred choice. Alternate techniques and their advantages are discussed further.

3.2.1. Negative Word line technique (NWL) [26, 29-32]

In this technique, the access transistors are biased in super cut-off region by applying a negative gate voltage instead of ground voltage during the hold mode as shown in Fig. 11. In SRAM cell based on NWL technique, an additional inverter (INV_P, INV_N) generates an appropriate WL signal to drive the access transistors. In the hold mode, negative gate-to-source voltage lowers the leakage current due to their exponential relationship in this region. Conversely, in read or write mode, the gate voltage is set to the nominal value as in conventional structure to achieve same performance.

3.2.2. Leakage biased bitline technique (LBB) [33-35]

The LBB technique is applicable to memories with hierarchical bitline architecture that comprises of sub-banks each having an individual set of local bitlines and pre-charge circuit. In this arrangement, only one sub-bank is activated at any given instance and other remain inactivated. A significant leakage current flows in the inactivated memory sub-banks. This current can be reduced by following LBB technique wherein pre-charge nMOS transistors with high threshold voltage is added in each sub-bank as shown in Fig. 12. It remains OFF in inactivated sub-banks while it conducts in the activated sub-bank. The bitlines BL and BLB are in floating state and are biased at random voltage levels depending on the leakage from the un-accessed cells. In other words, the leakage current self-bias the bitlines of an inactivated sub-bank. In the figure as can be seen (Fig. 12) the local BL2 of the inactivated sub-bank is disconnected from the pre-charge supply by keeping the nMOS pre-charge transistor OFF. By assuming the case that logic zero is stored (DB = '0') in all the cells then the Local BL2 will be biased to 0 V due to leakage current. Similarly, if all the cells store logic '1' then Local BL2 will be biased at V_{DD}. If half of the cells store '0' then the local BL2 will be biased somewhere at $V_{DD}/2$. Thus, the local bitline voltage is random and depends on the data stored in the column. It is observed that the technique results in memory array with lower leakage currents.

3.2.3. Comparative analysis

The SRAM based on the above techniques is implemented and the values of I_{BL_LEAK}, I_{READ}/I_{BL_LEAK} ratio and HSNM are measured. The simulations for the conditions specified earlier are performed and the corresponding results are plotted in Fig. 13 to Fig. 17. For NWL technique, a negative voltage generator of -80mV is used. For LBB technique, the theoretical concept is verified by simulating a column of SRAM cells storing equal number of ones and zeroes. In this condition the Local BL

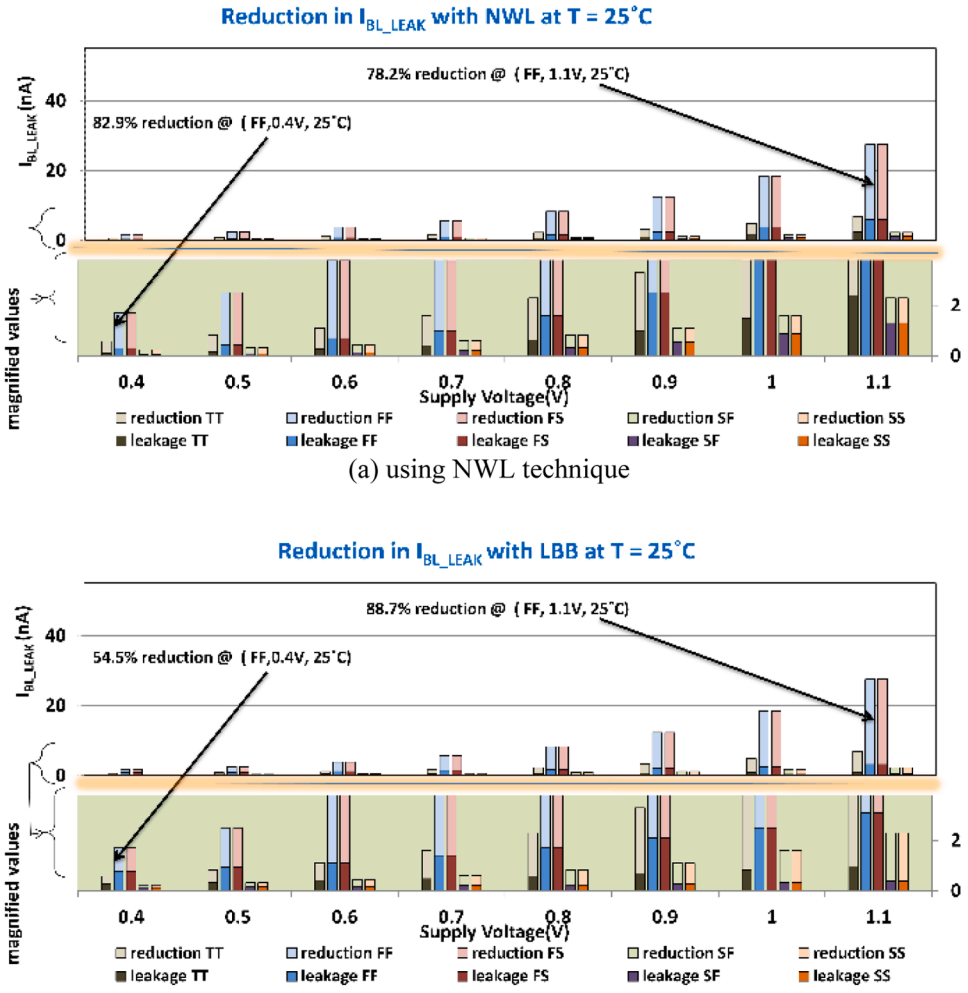


Fig. 14. Effect on Latch leakage current (I_{LATCH_LEAK}) at T = 25°C

get biased at $V_{DD}/2$.

3.2.3.1. Leakage reduction capabilities. At -40 ℃, the NWL technique achieves 69.2 % leakage reduction at 1.1 V which further increases to 80.9 % at 0.4 V as shown in Fig. 13(a). The opposite trend is observed with LBB technique, which offers 92.9 % reduction at 1.1 V and reduces to 62.2 % reduction at 0.4 V as depicted in Fig. 13(b). Thus, at lower temperatures, NWL technique gives best results at lower voltages while LBB technique gives best results at higher voltages. The plots for the I_{BL_LEAK} with PVT variations at 25 ℃ (Fig. 14) reveals the same trend as at -40° C. With further increase in temperature to 125 ℃, a decrease in percentage leakage reduction in I_{BL_LEAK} is observed (Fig. 15) for both the techniques over the entire voltage range. The percentage leakage reduction is 75 % and 47.1 % at 0.4 V at maximum leakage corner with NWL and LBB techniques respectively. It is worth mentioning that both the techniques provide same leakage performance at SS-SF and FF-FS corners at all temperatures.

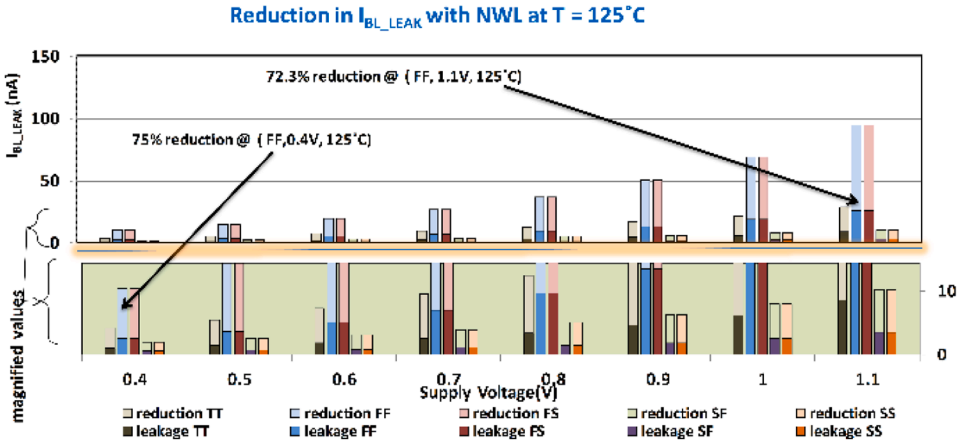
3.2.3.2. Impact on read performance. The read performance of NWL technique evaluated in terms of I_{READ}/I_{BL_LEAK} ratio improves as the temperature increases from -40 ℃ to 25 ℃ and then degrades with further increase in temperature as demonstrated in Fig. 16. The degradation at high temperatures is due to reduced I_{READ} values superimposed with inability of the technique to suppress leakages at these temperatures. Also, an improving trend as supply voltage scales from 1.1 V to 0.4 V is also inferred. Thus, it becomes clear from the results that the NWL technique offers the best read performance at moderate temperatures and lower voltages particularly in sub-threshold region. The LBB technique is not considered since it focuses on the leakage reduction in inactivated sub-banks.

3.2.3.3. Impact on hold stability. The HSNM results for the cells at 25°C are plotted in Fig. 17. The results (Fig. 17(a)) show that NWL technique has negligible effect on noise immunity of the cell especially at 0.4 V. However, a degradation in HSNM near 0.6 V at FF corner is observed. The HSNM results for LBB technique (Fig. 17(b)) show that like NWL technique it exhibits insignificant impact on HSNM, except for a slight degradation at SS corner near 0.4 V.

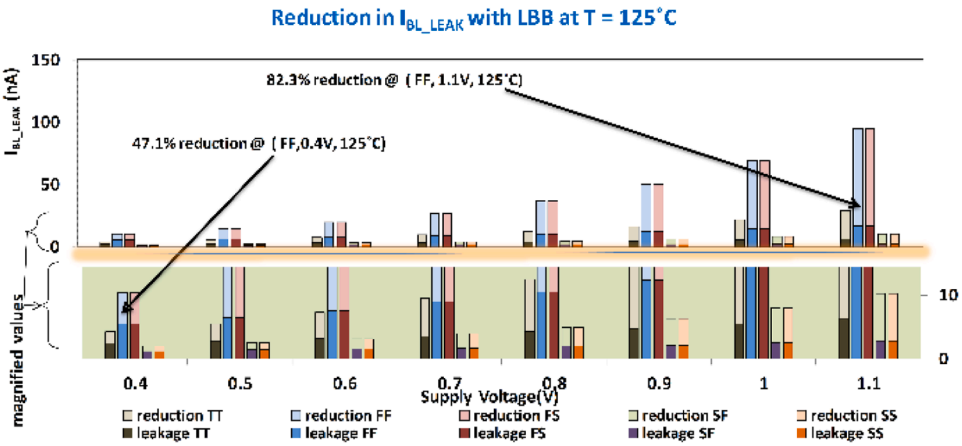
The performance of bitline leakage reduction techniques is outlined in Table 2 for quick comparison. The results indicate that the NWL technique is more suitable at low supply voltages and moderate temperatures whereas the LBB techniques is more appropriate for use at high supply voltages and low temperatures. Both, the techniques show improved tolerance to PVT-variations.

3.3. Read port leakage reduction techniques [3, 4, 6, 22, 25, 36-45]

The techniques covered in this sub-section are applicable to SRAM



(a) using NWL technique



(b) using LBB technique

Fig. 15. Effect on Latch leakage current (I_{LATCH_LEAK}) at T = 125°C

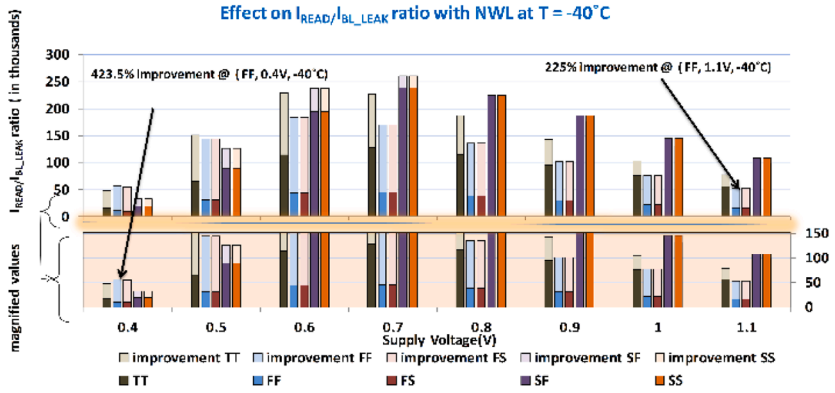
cells with an isolated/decoupled read port. The cells maintain a separate read port with a read bitline (RBL) common to all the cells in a column. During a read '0' operation, I_{READ} flows through the pre-charged RBL in the accessed cell to lower the RBL voltage whereas it remains high otherwise. Sense amplifier then senses the change in RBL voltage and produces the valid output. Thus, this arrangement provides SNM-free read operation as now the read current flows through the read port transistors. However, the simultaneous data dependent leakage flow (I_{RBL_LEAK}) in un-accessed cells may lead to the sensing of wrong data levels. Therefore, in the recent years the focus is shifted towards reducing the RBL leakage in read ports [4, 22, 36]. Various techniques, to address I_{RBL_LEAK}, are discussed below:

3.3.1. *Stack effect [6, 37-41]*

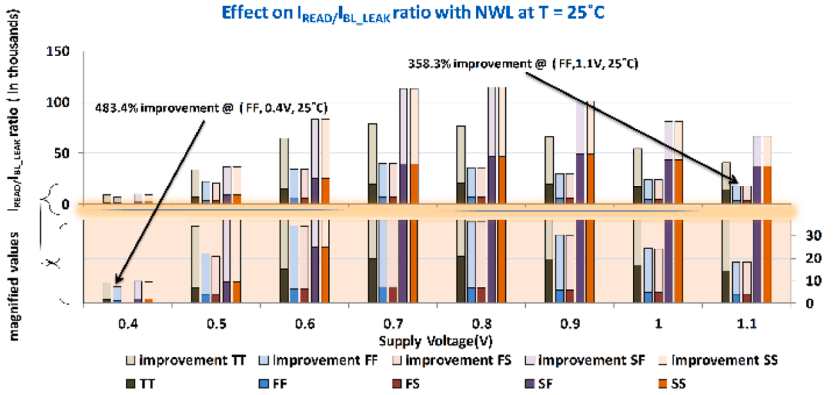
The Stack effect is one of the commonly used technique to reduce leakages in digital circuits. The approach is inherently static in nature as it provides the reverse biasing effect in both active mode and hold mode. There exists a variety of read ports that are designed on the basis of Stack effect. The read ports with a stack of two (RP1) and three transistors (RP2) are shown in Fig. 18. The read ports RP1 and RP2 are used in conventional 8T [6] and 9T [37] SRAM cells respectively. In RP2, the MOS transistor M1 of RP1 is split into M1a and M1b to enhance the Stacking effect. The additional OFF transistors in un-accessed cells increases the source-biasing effect hence causes lowering of I_{RBL_LEAK} values.

3.3.2. *Dynamic control of power rails (DCPR) [4]*

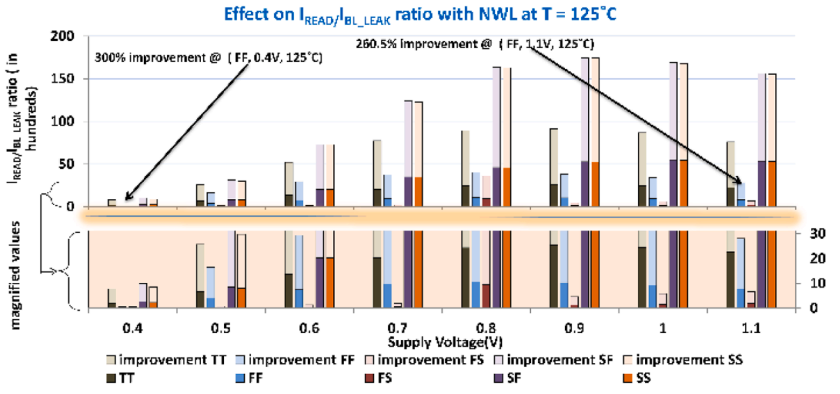
In this technique, the I_{RBL_LEAK} is suppressed by dynamically controlling the power rails of the read port. A virtually powered four transistor read port is shown in Fig. 19(a). It has a CMOS inverter (M1, M2) and transmission gate (M3, M4) whose power supply and control signal voltages are adjusted according to the operating mode of the cell. In the hold mode, the power rails (VVDD and VVSS) of the inverter and pre-charge level of RBL are kept at $V_{DD}/2$. The transmission gate control signals C and CB are set to '1' and '0' respectively thereby isolating the RBL from node 'X'. The voltage at node 'X' is raised to $V_{DD}/2$ via transistor M1 or M2 depending upon the value stored at internal node DB. This makes the potential of drain and source terminals of M3 and M4 equal thereby reducing I_{RBL_LEAK} significantly. During read mode, the power rails VVDD and VVSS are connected to V_{DD} and GND respectively. The control signal C is set to '0' whereas the complementary control signal CB is set to '1' thereby connecting the RBL (pre-charged to $V_{DD}/2$) to node 'X'. The use of transmission gate in place of single nMOS transistor improves the reading efficiency during read '1' operation.



(a) at T = -40 °C



(b) at T = 25 °C



(c) at T = 125 °C

Fig. 16. Effect on I_{READ}/I_{BL_LEAK} ratio using NWL technique

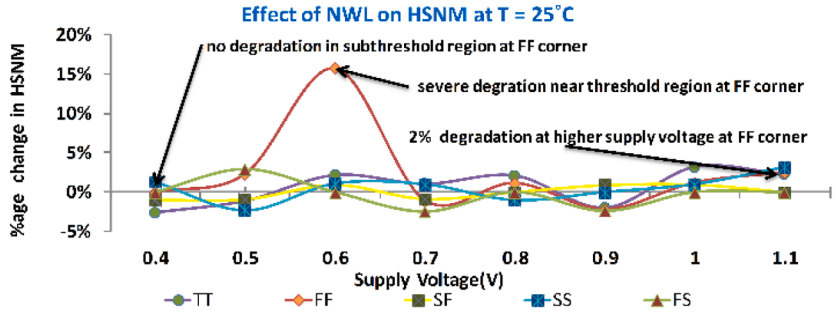
3.3.3. *Virtual cell ground (VCG) [6, 42-43]*

The VCG technique suggests (Fig. 19(b)) the use of virtual ground in SRAM cells. In this technique, the virtual ground terminal (RGND) of the read port is connected to high voltage during hold and write modes and to the ground during read mode. The effect of such an arrangement is the significant reduction in I_{RBL_LEAK} due to reduced drain to source voltage across the read port. This technique is versatile in the sense that it can be applied to the SRAM cells with either differential [43] or single ended read ports [6] very easily. The outcome is the significant reduction in I_{RBL_LEAK} and highly improved I_{READ}/I_{RBL_LEAK} ratio that results in tremendous increase in the number of memory cells that can be stacked per RBL.

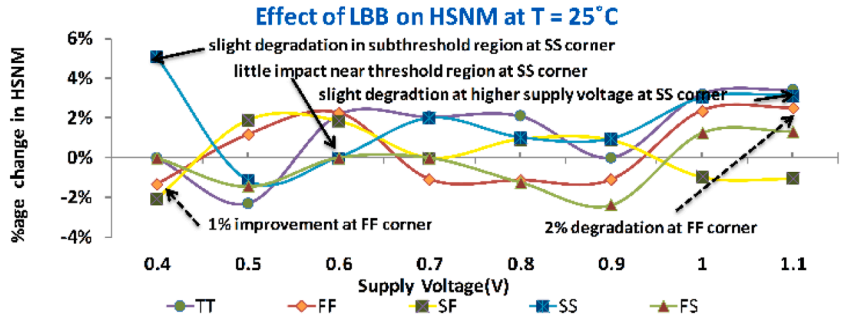
3.3.4. *Comparative analysis*

The comparison of all the techniques is done by simulating different SRAM cells with the conditions as previously mentioned (in sub-section 3.1). The results related to leakage reduction capability, read performance and hold stability are discussed in this sub-section.

3.3.4.1. Leakage reduction capabilities. At -40 ℃ (Fig. 20), it can be observed that stacking effect lowers I_{RBL_LEAK} by 93.1 % and 91.1 % at 1.1 V and 0.4 V respectively. The DCPR technique provides 96.1 %



(a) using NWL technique



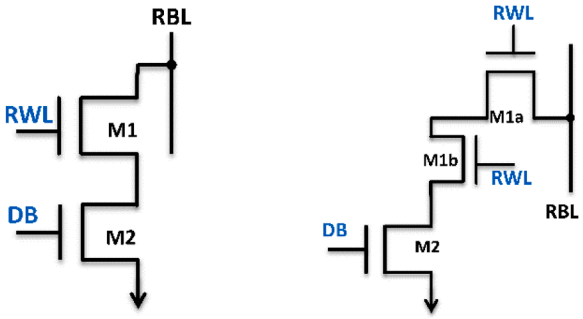
(b) using LBB technique

Fig. 17. Effect on Hold Static Noise Margin (HSNM) at T = 25°C

Table 2
Performance comparison of bitline leakage reduction techniques [25, 34-35]

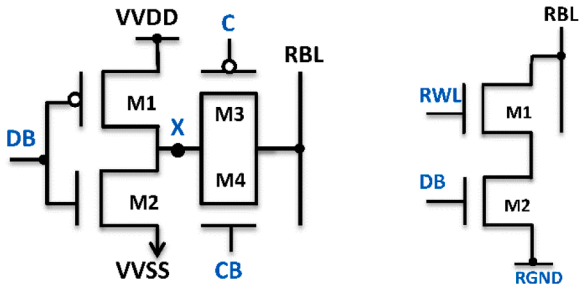
Techniques	NWL	LBB
Leakage	Overall good but more effective at low voltages	Overall good but more effective at high voltages
Read current to leakage current ratio	Improves at all voltages	No impact
Hold stability	Degradation at moderate voltages	Slight degradation at low voltages
PVT tolerant	Yes	Yes
Suitable temperatures	Moderate temperatures	Low temperatures
Suitable supply voltages	Low voltages	High voltages
Remarks	Good alternative to HVT MOS transistors at low voltages. Requirement of negative voltage generator.	Can be used to deactivate read paths in SRAM array with decoupled read port. Pre-charge phase is deferred when the corresponding part of memory is accessed for read/ write operation.

leakage reduction at 1.1 V and 85.7 % at 0.4 V while VCG technique reduces leakage by merely 4.8 % at 1.1 V and 72 % at 0.4 V. Therefore, it can be concluded that Stack effect and DCPR techniques are effective over the entire supply voltage range at lower temperatures while VCG technique works well at lower voltages only. As the temperature is increased to 25℃, Stack effect technique offer higher reduction at high voltages whereas it decreases at lower voltages as shown in Fig. 21(a). The DCPR technique (Fig. 21(b)) on the other hand provides the impeccable performance irrespective of changes in temperature. The performance of VCG technique, improves with increase in temperature as the percentage leakage reduction increases to 30.3% at 1.1 V and 95.4 % at 0.4 V as reflected in Fig. 21(c). At 125℃, a similar trend is observed as at 25℃ for all the techniques as shown in Fig. 22. Therefore, it can be



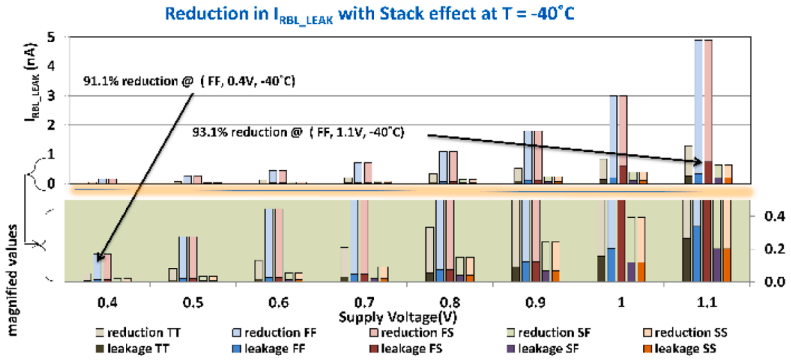
(a) RP1 [6] (b) RP2 [37]

Fig. 18. Isolated read ports of SRAM cells with Stack effect

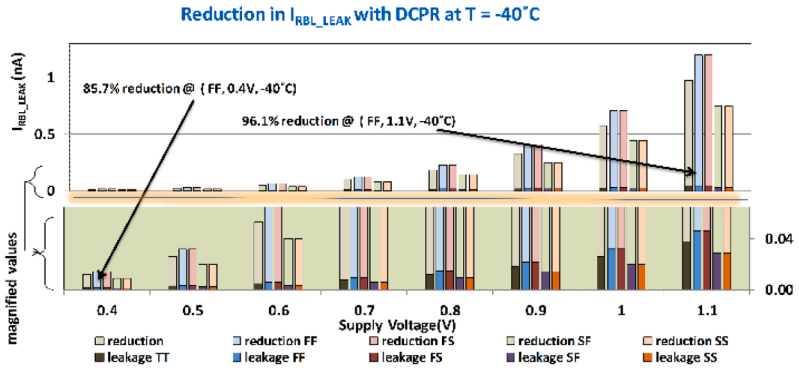


(a) using DCPR technique [4] (b) using VCG technique [31]

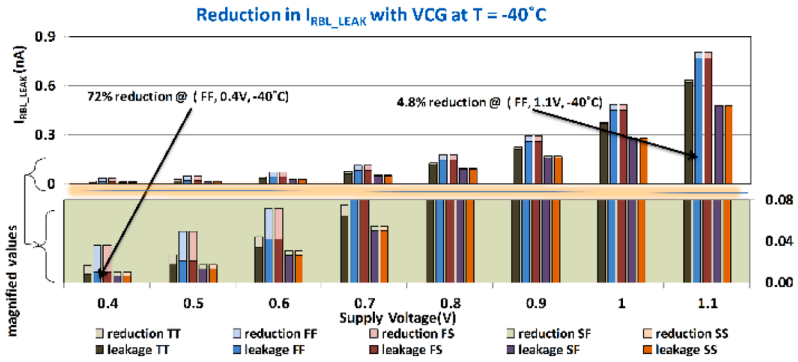
Fig. 19. Isolated read port of SRAM cells



(a) using Stack Effect



(b) using DCPR technique

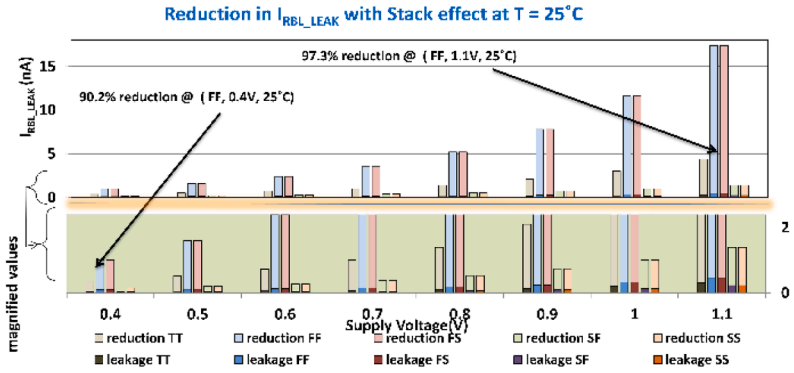


(c) using VCG technique

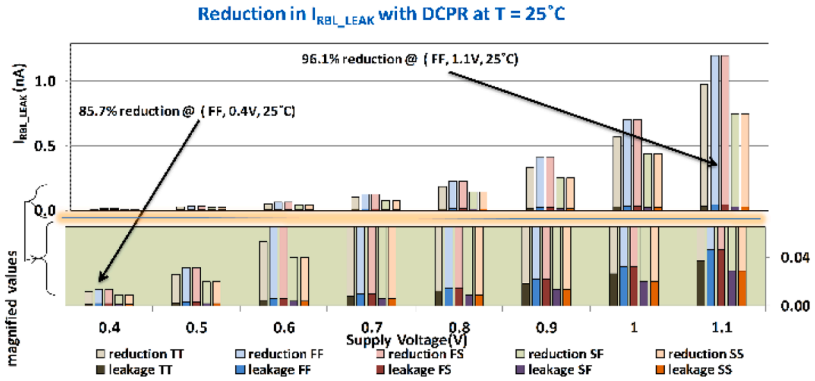
Fig. 20. Effect on Read bitline leakage current (I_{RBL_LEAK}) at T = -40°C

concluded that both Stack effect and DCPR techniques has extremely good leakage reduction capabilities over the entire supply voltage range while VCG technique is good at low voltages particularly in sub-threshold region. It can be seen from the leakage reduction results plotted in Fig. 20 to Fig. 22 that all the three techniques show same leakage reduction capabilities at FF-FS and SF-SS corners at all temperatures particularly below 0.9 V.

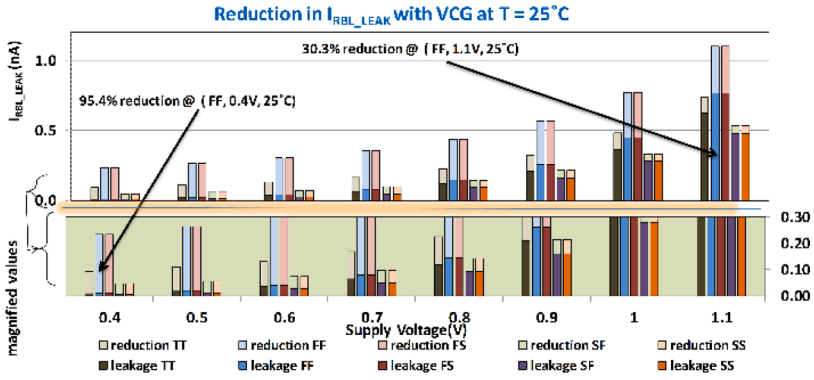
3.3.4.2. Impact on read performance. The impact of Stack effect, DCPR and VCG techniques on read performance is captured in terms of I_{READ}/I_{RBL_LEAK} ratio and is plotted for -40 ℃, 25 ℃ and 125 ℃ in Fig. 23, Fig. 24 and Fig. 25 respectively. It can be observed for the read port design, based on Stack effect technique, both I_{RBL_LEAK} and I_{READ} values are reduced at all process corners, temperatures and supply voltages. However, due to more effect on leakages, the improvement in I_{READ}/I_{RBL_LEAK} ratio is achieved irrespective of operating conditions. Since the other two techniques are dynamic in nature so there is no effect on read current values but, due to reduced leakages, there are subsequent changes in I_{READ}/I_{RBL_LEAK} ratio of read port. At -40 ℃, the percentage improvement in I_{READ}/I_{RBL_LEAK} ratio for Stack effect technique is 477.9 % at 1.1 V and 292.1 % at 0.4 V. With the DCPR technique, improvement in I_{READ}/I_{RBL_LEAK} ratio is 2491.8 % and 600 % at 1.1 V and 0.4 V respectively. Similarly, in case of VCG technique it is merely 5.1 % at 1.1 V and 257.4 % at 0.4 V. It is to be noted that though all the techniques achieve significant improvement in I_{READ}/I_{RBL_LEAK} ratio but out of all three techniques, DCPR technique gives the highest improvement over the entire voltage range. At 25 ℃ (Fig. 24), the percentage improvement of Stack effect in I_{READ}/I_{RBL_LEAK} ratio increases to 1335.6 % at 1.1 V and



(a) using Stack Effect



(b) using DCPR technique

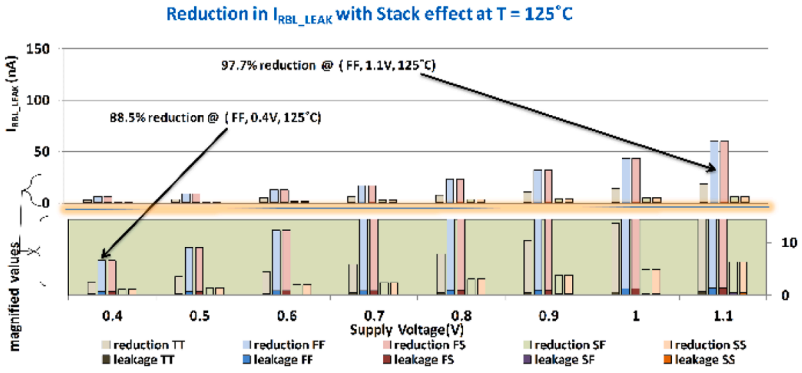


(c) using VCG technique

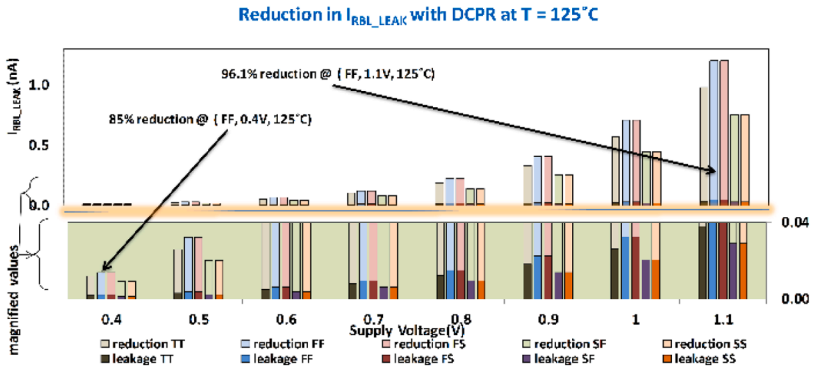
Fig. 21. Effect on Read bitline leakage current (I_{RBL_LEAK}) at T = 25°C

reduces to 242.7 % at 0.4 V due to its strong leakage reduction capabilities at high voltages and reduced read current values at lower voltages. The performance of DCPR technique remains unchanged with increase in temperature while that of VCG technique improves. For VCG technique, I_{READ}/I_{RBL_LEAK} ratio increases to 43.5 % and 2076.9 % at 1.1 V and 0.4 V respectively. As the temperature increases from 25 ℃ to 125 ℃ (Fig. 25), the percentage improvement in I_{READ}/I_{RBL_LEAK} ratio for Stack effect technique further rises to 1439.1 % at 1.1 V and drops to 196.2 % at 0.4 V due to the reasons mentioned earlier. However, as the result shows, the percentage improvement remains significant throughout the voltage range irrespective of temperatures even though it shows better results at higher temperature. This makes the technique, preferred choice at extremely higher temperatures for achieving high performance with leakage resilience. The DCPR technique keeps on achieving the same performance irrespective of changes in temperature. For VCG technique, significant improvement of more than four orders of magnitude is observed at 0.4 V. Thus, it can be concluded that with extremely good results at lower voltages, the VCG technique is best for use in sub-threshold region at higher temperatures for low power SRAM designs.

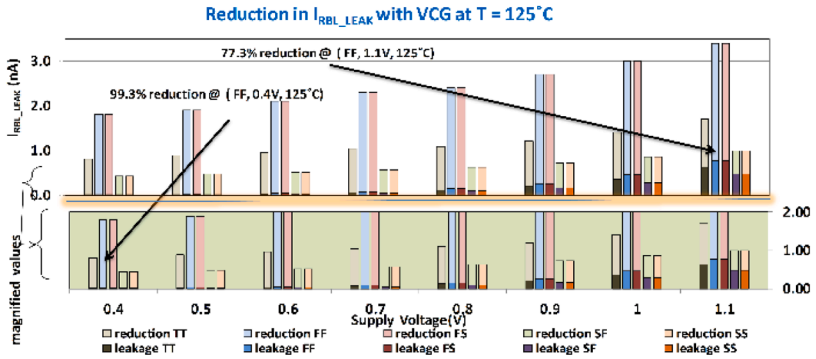
3.3.4.3. Impact on hold stability. The hold stability expressed in terms of HSNM at 25 ℃ for all the techniques is captured and plotted in Fig. 26. It is observed that none of the techniques namely Stack effect, DCPR and



(a) using Stack Effect



(b) using DCPR technique



(c) using VCG technique

Fig. 22. Effect on Read bitline leakage current (I_{RBL_LEAK}) at T = 125°C

VCG show any effect on HSNM values and the values vary only as corner and/or supply voltage change. Thus, the hold stability is greatly affected by increase in variability and the factors which are internal to the latch such as pull-up, pull-down ratio etc. [44, 45].

The performance of read port leakage reduction techniques are tabulated on the basis of leakage, read performance and stability in Table 3 for quick comparison. The results indicate that the stack effect and DCPR techniques are more effective at high supply voltages and low-to-moderate temperatures in contrast to VCG technique which is more suitable at low supply voltages and high temperatures. It is also noted that all the techniques show improved tolerance to PVT-variations.

4. Technology node comparison

The performance evaluation of all the leakage reduction techniques for SRAM cells is carried out at different technology nodes in sub-threshold region and 25°C. The SRAM cell performance parameters namely read stability, write ability and minimum retention voltage are also considered for completeness. The read stability and write ability are measured in terms of Read Static Noise Margin (RSNM) and Write Margin (WM) respectively. The results with and without techniques are

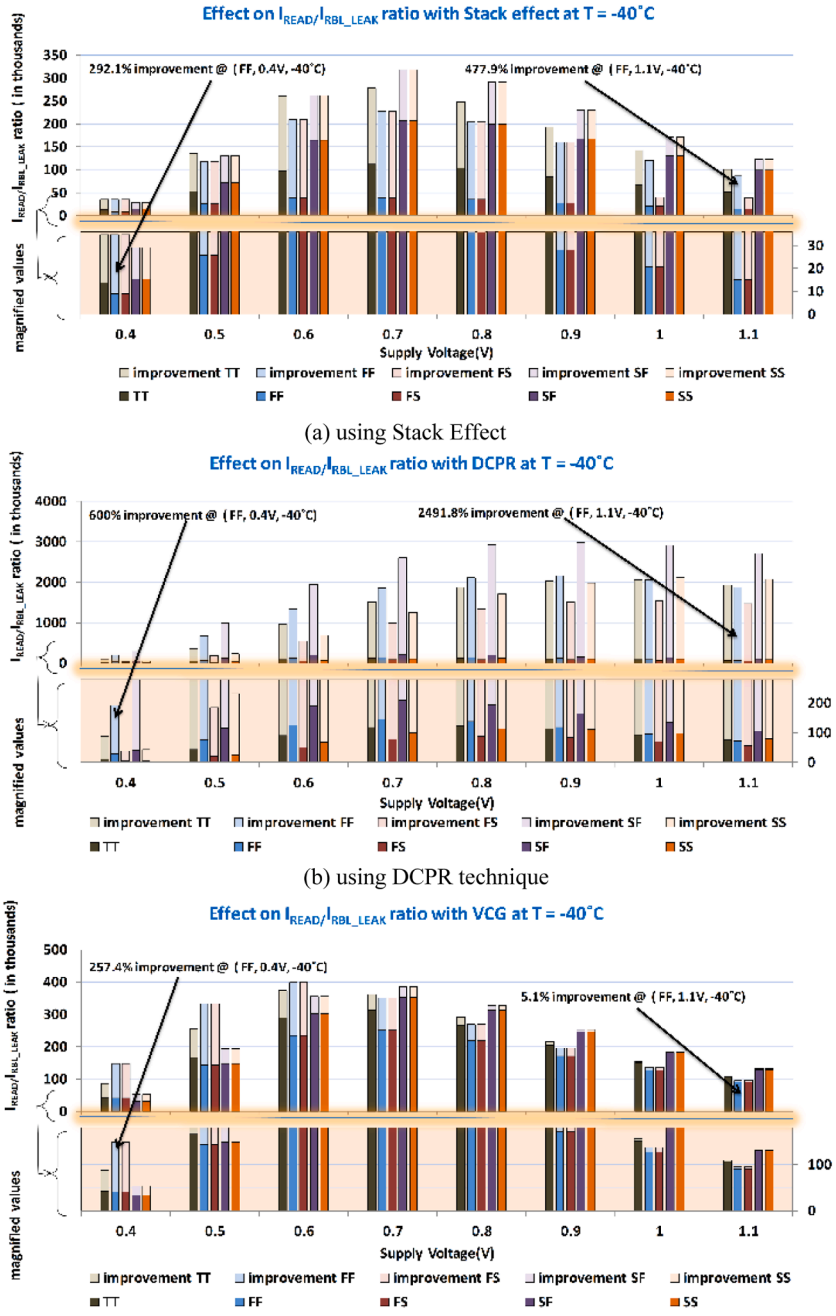


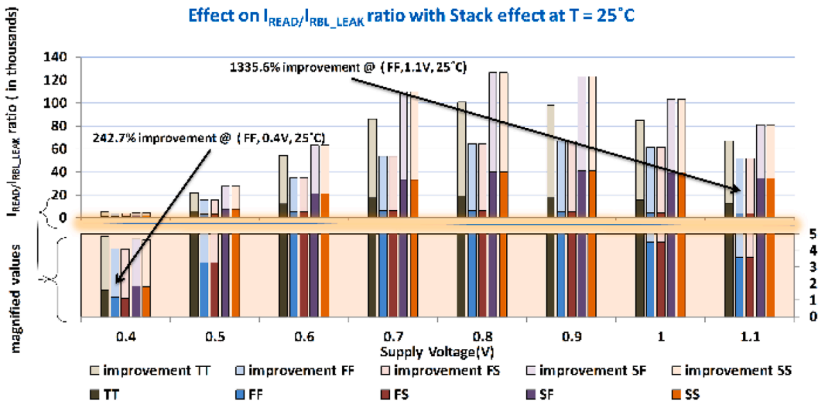
Fig. 23. Effect on I_{READ}/I_{RBL_LEAK} ratio at T = -40°C

captured and changes observed in percentages at different technology nodes are summarized in Table 4, Table 5 and Table 6.

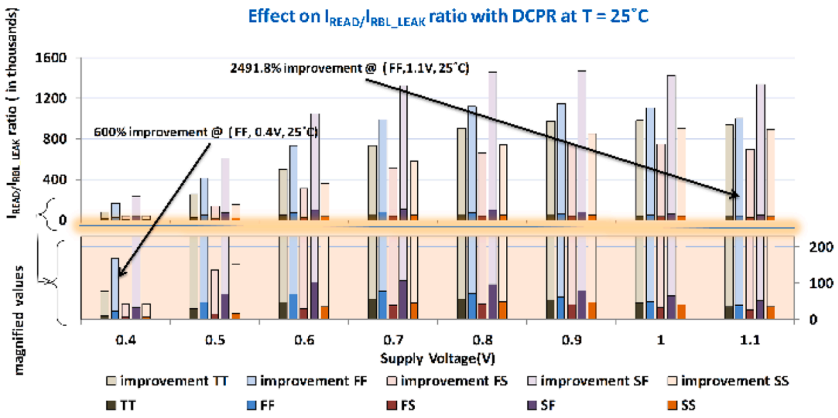
The results for latch leakage reduction techniques (Table 4) indicate that the efficacy of MTCMOS technique in suppressing leakage current remains significant across the technology nodes in comparison to SB and DM techniques. It is also noted that MTCMOS technique being static in nature has degrading effect on read stability of the SRAM cell. However, a significant improvement in write ability and minimum retention voltage is also observed. Additionally, a significant improvement in minimum retention voltage and severe degradation in hold stability are also observed with SB and DM techniques respectively.

Further, from the results in Table 5, it is observed that for the techniques aimed at reducing bitline leakages, the effectiveness of NWL technique remains high in contrast to LBB technique with technology scaling. In addition, a significant improvement in read current to leakage current ratio is noted with NWL technique at the cost of minimum retention voltage. There is no impact of LBB technique on read/write parameters as it focuses on the leakage reduction in inactivated sub-banks.

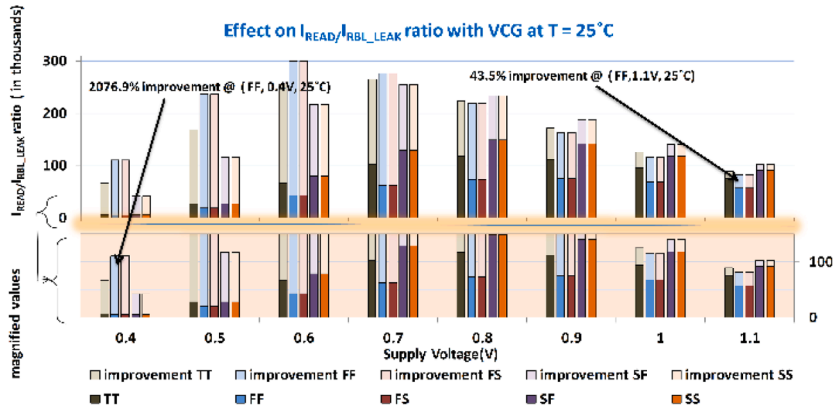
Among the techniques intended for reducing the read port leakages (Table 6), all the three techniques, namely Stack effect, DCPR and VCG techniques show impeccable leakage reduction capability across the technology nodes. However, VCG technique retains significantly high leakage reduction percentage compared to the other two. In addition, no impact of any technique on cell stability, write ability and minimum retention voltage is observed as all the techniques are applied on the



(a) using Stack Effect

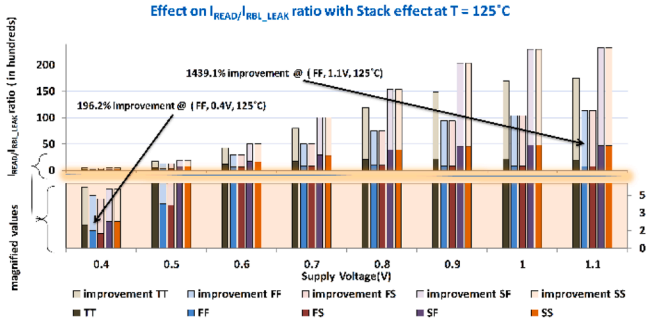


(b) using DCPR technique

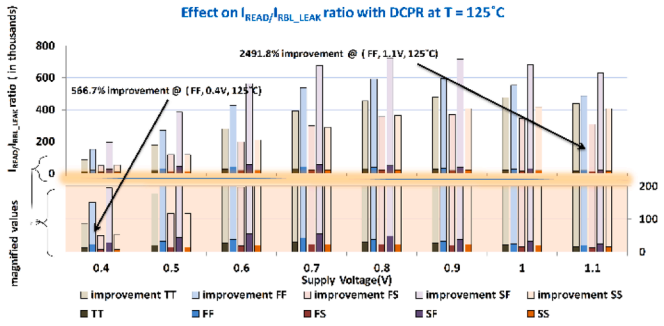


(c) using VCG technique

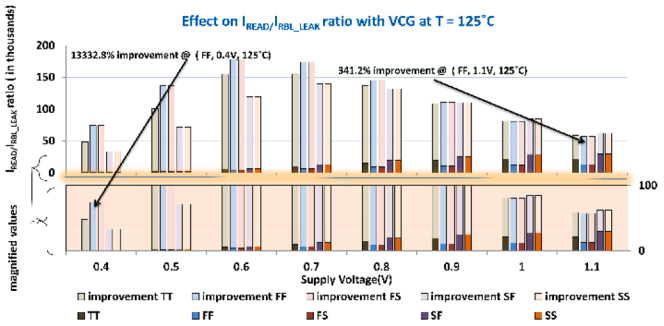
Fig. 24. Effect on I_{READ}/I_{RBL_LEAK} ratio at T = 25°C



(a) using Stack Effect



(b) using DCPR technique



(c) using VCG technique

Fig. 25. Effect on I_{READ}/I_{RBL_LEAK} ratio at T = 125°C

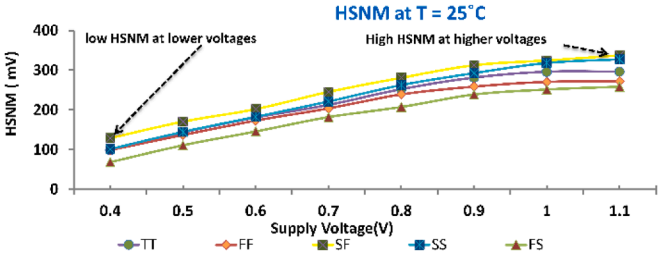


Fig. 26. Hold Static Noise Margin (HSNM) at T = 25°C for Stack effect, DCPR and VCG techniques

Table 3
Performance comparison of read port leakage reduction techniques [25, 39, 42-45]

Techniques	STACK EFFECT	DCPR	VCG
Leakage	Overall good but more effective at high voltages	Overall good but more effective at high voltages	Effective at low voltages
Read current to leakage current ratio	More improvement at high voltages	More improvement at high voltages	More improvement at low voltages
Hold stability	No effect	No effect	No effect
PVT tolerant	Yes	Yes	Yes
Suitable temperatures	Low to moderate temperatures	-40°C to 125°C	High temperatures
Suitable supply voltages	High voltages	High voltages	low voltages
Remarks	Degraded read current values at high temperatures. Use of extra transistors.	Requirement of multiple supply rails.	Requirement of multiple ground rails.

SRAM cell with isolated read port.

5. Conclusion

The techniques with leakage reduction capabilities should bear least effect on the performance parameters of the circuit. So, in the paper, the impact of each technique on read performance and stability has been estimated in addition to the percentage reduction in leakages. It is observed that for reducing latch leakages, MTCMOS is more suitable at high voltages and temperatures whereas DM technique is more suitable at high voltages and over the temperature range from -40°C to 125°C. The SB technique is suitable over the entire temperature (-40°C to 125°C) and voltages range (0.4 V to 1.1 V). It is also found that all the techniques show strong impact of process variations on the leakage reduction capabilities.

To suppress the bitline leakages, results for both NWL and LBB techniques show significant leakage reduction across temperature and supply voltage range. The suitable operating conditions for NWL technique are moderate temperatures and low voltages. The LBB technique reduces the leakage through memory sub-bank significantly at low temperatures and high voltages independent of the values stored in the cell. Both, the techniques show lesser impact of process variations on the leakage reduction capabilities.

For reducing read port leakages, the results show that the Stack effect gives extremely high percentage reduction in leakages across all voltages and temperatures but at higher temperature there exists a trade-off between read performance and leakage reduction capabilities making this technique suitable for low to moderate temperatures. The DCPR technique, on the other hand, achieves extremely good results at high voltages and over a temperature range from -40°C to 125°C while VCG technique is suitable for use at high temperatures and low voltages. It is also observed that all the three techniques Stack effect, DCPR and VCG techniques possess tolerance for leakages against process variations.

On performing the technology comparison on the leakage reduction techniques, it is observed that the efficacy of the techniques in suppressing the leakage current changes with technology node. In addition, all the techniques bear either positive or no impact on read-write performance of the SRAM cell except MTCMOS technique. The MTCMOS technique, provides a significant improvement in write ability and minimum retention voltage of the SRAM cell but at the cost of read stability.

Declaration of Competing Interest

There is no conflict of interest.

Table 4
Comparison of latch leakage reduction techniques at different technology nodes [15] [24-28]

Techniques	MTCMOS			SB			DM		
	90 nm	65 nm	32 nm	90 nm	65 nm	32 nm	90 nm	65 nm	32 nm
Leakage	-97.0 %	-94.4 %	-90.5 %	-75.9 %	-74.1 %	-45.2 %	-12.8 %	-17.1 %	-54.1 %
Read current	-93.8 %	-99.9 %	-99.8 %	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect
Hold stability	-3.8 %	-5 %	-5.3 %	+6 %	+5 %	+4 %	-75 %	-78 %	-83 %
Read stability	-74.6 %	-61.9 %	-99.9 %	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect
Write ability	+106.3 %	+102.7 %	+83.2 %	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect
Minimum retention voltage	-70.2 %	-62.1 %	-45.3 %	-37.9 %	-31.8 %	-20.9 %	Negligible Effect	Negligible Effect	Negligible effect

Table 5
Comparison of bitline leakage reduction techniques at different technology nodes [25, 34-35]

Techniques	NWL			LBB		
	90 nm	65 nm	32 nm	90 nm	65 nm	32 nm
Leakage	-87.6 %	-85.2 %	-81.1 %	-27.7 %	-36.5 %	-50.5 %
Read current to leakage current ratio	+703.8 %	+577.8 %	+425.7 %	No Effect	No Effect	No Effect
Hold stability	Negligible Effect	Negligible Effect	Negligible effect	No Effect	No Effect	No Effect
Read stability	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect
Write ability	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect
Minimum retention voltage	+15.7 %	+12.9 %	+12.6 %	No Effect	No Effect	No Effect

Table 6
Comparison of read port leakage reduction techniques at different technology nodes [25, 39, 42-45]

Techniques	STACK EFFECT			DCPR			VCG		
	90 nm	65 nm	32 nm	90 nm	65 nm	32 nm	90 nm	65 nm	32 nm
Leakage	-81 %	-84.6 %	-89.2 %	-84.0 %	-87.2 %	-85.8 %	-98.7 %	-95.9 %	-90.8 %
Read current to leakage current ratio	+115.5 %	+145.5 %	+206.2 %	+525 %	+680 %	+605.9 %	+7800 %	+2333 %	+984.9 %
Hold stability	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect
Read stability	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect
Write ability	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect
Minimum retention voltage	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect	No Effect

References

[1] ITRS: International technology road map for semiconductors, test and test equipments. http://www.itrs2.net/.
[2] S. Thompson, P. Packan, M. Bohr, MOS Scaling, Transistor Challenges for the 21st Century, Intel Technology Journal Q3 (1998).
[3] G. Sery, S. Borkar, V. De, Life is CMOS: Why chase the life after?, in: Proc. - Des. Autom. Conf., 2002.
[4] N. Maroof, B.S. Kong, 10T SRAM Using Half-VDD Precharge and Row-Wise Dynamically Powered Read Port for Low Switching Power and Ultralow RBL Leakage, IEEE Trans. Very Large Scale Integr. Syst. (2017).
[5] A. Agarwal, S. Mukhopadhyay, A. Raychowdhury, K. Roy, C.H. Kim, Leakage power analysis and reduction for nanoscale circuits, IEEE Micro. (2006).
[6] L. Chang, R.K. Montoye, Y. Nakamura, K.A. Batson, R.J. Eickemeyer, R.H. Dennard, W. Haensch, D. Jamsek, An 8T-SRAM for variability tolerance and low-voltage operation in high-performance caches, in: IEEE J. Solid-State Circuits, 2008.
[7] J.J. Kim, A. Bansal, R. Rao, S.H. Lo, C. Te Chuang, Relaxing conflict between read stability and writability in 6T SRAM cell using asymmetric transistors, IEEE Electron Device Lett. (2009).
[8] T.P. Haraszti, CMOS Memory Circuits, Kluwer Academic Publishers, 2000, p. 522.
[9] C.C. Wang, P.M. Lee, K.L. Chen, 6-T SRAM using dual threshold voltage transistors and low-power quenchers, in: Proc. IEEE Int. Conf. Electron. Circuits, Syst., 2002.
[10] B. Amelifard, F. Fallah, M. Pedram, Low-leakage SRAM design with dual V/sub t/ transistors, in: Proc. - Int. Symp. Qual. Electron. Des. ISQED, 2006.
[11] Siva G Narendra, Anantha Chandrakasan, Leakage in Nanometer CMOS Technologies, Springer, 2006, pp. 8–172.
[12] S.A. Tawfik, V. Kursun, Low power and robust 7T dual-Vt SRAM circuit, Proc. - IEEE Int. Symp. Circuits Syst. (2008) 1452–1455.
[13] R. Faraji, H.R. Naji, M. Rahimi-Nezhad, M. Arabnejhad, New SRAM design using body bias technique for low-power and high-speed applications, Int. J. Circuit Theory Appl. (2014).
[14] R. Gonzalez, B.M. Gordon, M.A. Horowitz, Supply and threshold voltage scaling for low power CMOS, IEEE J. Solid-State Circuits. (1997).
[15] C.H. Kim, J.J. Kim, S. Mukhopadhyay, K. Roy, A forward body-biased low-leakage SRAM cache: Device, circuit and architecture considerations, IEEE Trans. Very Large Scale Integr. Syst. (2005).
[16] R. Suthar, K.S. Pande, N.S. Murty, Leakage reduction in DT8T SRAM cell using body biasing technique, in: Proc. - 2017 IEEE Int. Symp. Nanoelectron. Inf. Syst. INIS 2017, 2018.
[17] N.S. Kim, K. Flautner, D. Blaauw, T. Mudge, Circuit and Microarchitectural Techniques for Reducing Cache Leakage Power, in: IEEE Trans. Very Large Scale Integr. Syst., 2004.
[18] L.T. Clark, R. Patel, T.S. Beatty, Managing Standby and Active Mode Leakage Power in Deep Sub-micron Design, in: Proc. Int. Symp. Low Power Electron. Des., 2004.
[19] T.D. Burd, T.A. Pering, A.J. Stratakos, R.W. Brodersen, Dynamic voltage scaled microprocessor system, IEEE J. Solid-State Circuits. (2000).
[20] N. Edri, S. Fraiman, A. Teman, A. Fish, Data retention voltage detection for minimizing the standby power of SRAM arrays, in: 2012 IEEE 27th Conv. Electr. Electron. Eng. Isr. IEEEI 2012, 2012.
[21] L.T. Clark, M. Morrow, W. Brown, Reverse-body bias and supply collapse for low effective standby power, IEEE Trans. Very Large Scale Integr. Syst. (2004).
[22] G. Pasandi, S.M. Fakhraie, A 256-kb 9T Near-Threshold SRAM with 1k Cells per Bitline and Enhanced Write and Read Operations, IEEE Trans. Very Large Scale Integr. Syst. (2015).
[23] N. H.E.Weste, D.M. Harris, CMOS VLSI Design: A Circuits and Systems Perspective (4th Edition), 2010.
[24] B. Wang, T.Q. Nguyen, A.T. Do, J. Zhou, M. Je, T.T.H. Kim, Design of an ultra-low voltage 9T SRAM with equalized bitline leakage and CAM-assisted energy efficiency improvement, IEEE Trans. Circuits Syst. I Regul. Pap. 62 (2015) 441–448.
[25] K. Roy, S. Mukhopadhyay, H. Mahmoodi-Meimand, Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits, Proc. IEEE. (2003).
[26] Y.C. Lai, S.Y. Huang, X-calibration: A technique for combating excessive bitline leakage current in nanometer SRAM designs, IEEE J. Solid-State Circuits. (2008).
[27] J. Wu, Y. Zhang, M. Zukerman, E.K.N. Yung, Energy-efficient base-stations sleep-mode techniques in green cellular networks: A survey, IEEE Commun. Surv. Tutorials. (2015).
[28] I. Ashraf, F. Boccardi, L. Ho, SLEEP mode techniques for small cell deployments, IEEE Commun. Mag. (2011).
[29] Yang et. al., Negative word line driver for semiconductor memories, United States Patent, Sep. 9, 2014.
[30] Z. Liu, V. Kursun, Characterization of a novel nine-transistor SRAM cell, IEEE Trans. Very Large Scale Integr. Syst. (2008).
[31] H. Kawaguchi, K. Nose, T. Sakurai, Super cut-off CMOS (SCCMOS) scheme for 0.5-V supply voltage with picoampere stand-by current, IEEE J. Solid-State Circuits. 35 (2000) 1498–1501.
[32] D. Kilani, B. Mohammad, M. Alhawari, H. Saleh, M. Ismail, Power Management for Wearable Electronic Devices, Springer Nature, 2020.

[33] S. Heo, K. Barr, M. Hampton, K. Asanović, Dynamic fine-grain leakage reduction using leakage-biased bitlines, in: Conf. Proc. - Annu. Int. Symp. Comput. Archit. ISCA, 2002: pp. 137–147.
[34] A. Karandikar, K.K. Parhi, Low power SRAM design using hierarchical divided bit-line approach, in: Proc. - IEEE Int. Conf. Comput. Des. VLSI Comput. Process., 1998.
[35] B. D. Yang, L.S. Kim, A low-power SRAM using hierarchical bit line and local sense amplifiers, IEEE J. Solid-State Circuits. (2005).
[36] K. Agawa, H. Hara, T. Takayanagi, T. Kuroda, A bitline leakage compensation scheme for low-voltage SRAMs, IEEE J. Solid-State Circuits. (2001).
[37] S.A. Verkila, S.K. Bondada, B.S. Amrutur, A 100MHz to IGHz, 0.35V to 1.5V supply 256 × 64 SRAM block using symmetrized 9T SRAM cell with controlled read, in: Proc. IEEE Int. Freq. Control Symp. Expo., 2008.
[38] S. Narendra, S. Borkar, V. De, D. Antoniadis, A. Chandrakasan, Scaling of stack effect and its application for leakage reduction, in: Proc. Int. Symp. Low Power Electron. Des. Dig. Tech. Pap., 2001.
[39] S. Ahmad, M.K. Gupta, N. Alam, M. Hasan, Single-Ended Schmitt-Trigger-Based Robust Low-Power SRAM Cell, IEEE Trans. Very Large Scale Integr. Syst. (2016).
[40] J.P. Kulkarni, K. Roy, A. We, S. St, Ultralow-Voltage Process-Variation-Tolerant Schmitt-Trigger-Based SRAM Design 20 (2012) 319–332.
[41] A. Saxena, A. Shrivastava, S. Akashe, Estimation of high performance in Schmitt triggers with stacking power-gating techniques in 45 nm CMOS technology, Int. J. Commun. Syst. (2014).
[42] A. Sil, S. Bakkamanthala, S. Karlapudi, M. Bayoumi, Highly stable, dual-port, sub-threshold 7T SRAM cell for ultra-low power application, 2012 IEEE 10th Int. New Circuits Syst. Conf. NEWCAS 2012. (2012) 493–496.
[43] T. Song, S. Kim, K. Lim, J. Laskar, Fully-gated ground 10T-SRAM bitcell in 45nm SOI technology, Electron. Lett. (2010).
[44] J.P. Kulkarni, K. Kim, K. Roy, A 160 mV robust schmitt trigger based sub threshold SRAM, IEEE J. Solid-State Circuits. 42 (2007) 2303–2313.
[45] N. Rahman, B.P. Singh, Static-Noise-Margin Analysis of Conventional 6T SRAM Cell at 45nm Technology, Int. J. Comput. Appl. (2013).



Monica Gupta received M.E. degree in Electronics and Communication Engineering from Delhi College of Engineering, New Delhi in the year 2006. She is currently pursuing Ph. D degree in Electronics and Communication Engineering from Delhi Technological University, New Delhi (Formerly, DCE). She is working as an Assistant Professor with Bharati Vidyapeeth's College of Engineering, New Delhi. Her research interests include Low power memory design.



Kirti Gupta received B.Tech. in Electronics and Communication Engineering from Indira Gandhi Institute of Technology, Delhi in 2002, M. Tech. in Information Technology from School of Information Technology in 2006. She received her Ph.D. in Electronics and Communication Engineering from Delhi Technological University, in 2016. Since 2002, she is with Bharati Vidyapeeth's College of Engineering, New Delhi and is presently serving as Professor in the same institute. A life member of ISTE, and senior member of IEEE, she has published more than 100 research papers in international journals and conferences. She has recently co-authored a book with Springer Nature. Her teaching and research interest is in digital VLSI design.



Neeta Pandey received her M.E. in Microelectronics from Birla Institute of Technology and Sciences, Pilani in 1991 and Ph.D. from Guru Gobind Singh Indraprastha University, Delhi in 2009. She has served in Central Electronics Engineering Research Institute, Pilani, Indian Institute of Technology, Delhi, Priyadarshini College of Computer Science, Noida and Bharati Vidyapeeth's College of Engineering, Delhi in various capacities. At present, she is a professor in the ECE department, Delhi Technological University. Her teaching and research interests include analog and digital VLSI design. A life member of ISTE, and senior member of IEEE, USA, she has co-authored over 200 papers in international, national journals of repute and confer ences. Recently, she has co-authored a book with Springer Nature.

2021 IEEE/ACM 18th International Conference on Mining Software Repositories (MSR)

Comparative Study of Feature Reduction Techniques in Software Change Prediction

Ruchika Malhotra
Department of Software Engineering
Delhi Technological University
Bawana Road, Delhi, India
ruchikamalhotra@dtu.ac.in

Ritvik Kapoor
Department of Software Engineering
Delhi Technological University
Bawana Road, Delhi, India
ritvikush953@gmail.com

Deepti Aggarwal
Department of Software Engineering
Delhi Technological University
Bawana Road, Delhi, India
deeptiagg16@gmail.com

Priya Garg
Department of Software Engineering
Delhi Technological University
Bawana Road, Delhi, India
gargpriya1304@gmail.com

***Abstract*— Software change prediction (SCP) is the process of identifying change-prone software classes using various structural and quality metrics by developing predictive techniques. The previous studies done in this field strongly confer the correlation between the quality of metrics and the performance of such SCP models. Past SCP studies have also applied different feature reduction (FR) techniques to address issues of high dimensionality, feature irrelevance, and feature repetition. Due to the vast variety of metric suites and FR techniques applied in SCP, there is a need to analyze and compare them. It will help in identifying the most crucial features and the most effective FR techniques. So, in this research, we conduct experiments to compare and contrast 60 Object-Oriented plus 26 Graph-based metrics and 11 state-of-the-art FR techniques previously employed for SCP over a range of 6 Java projects and 3 diverse classifiers. The AUC-ROC measures and statistical tests over experimental SCP models indicate that FR techniques are effective in SCP. Also, there exist significant differences in the performance of the different FR techniques. Furthermore, from this extensive experimentation, we were able to identify a set of the most effective FR techniques and the most crucial metrics which can be used to build effective SCP models.**

Keywords—Software change prediction, Feature reduction, Object-Oriented metrics, Graph-based metrics, Machine Learning

I. INTRODUCTION

Iterative software development, functionality updates, and extensive software maintenance support are becoming a standard in today's dynamic software industry. Accommodating customer's demands, code refactoring, changes in technology, and fault correction are some of the reasons which contribute to code changes leading to newer versions and releases of the software [1]. These change-related activities can be costly and time-consuming for large and long-running software projects. Hence, industries are nowadays employing Software Change Prediction (SCP) to identify the classes that are most prone to changes in the subsequent releases of the software [2, 3, 4]. Identification of such change-prone classes aids in careful planning, easy maintenance, and proper management of the software projects. It helps the project managers to focus their limited resources on such change-prone classes in advance. It ultimately leads to the development of high-quality software that is maintained and updated in the scheduled time [2].

Recently, various Machine Learning (ML) techniques have been used to predict the change-proneness of the classes [2-5]. These techniques use a variety of independent features for prediction like OO metrics [5], class dependency graph-based metrics [6], class evolution-based metrics [7], developer-related factors [8], etc. Many studies often combine different metric suites to obtain higher performance in SCP [9]. However, this increases the dimensionality of the dataset and inadvertently introduces many irrelevant and redundant features [10]. Working with such high-dimensional datasets brings several challenges like high-computation costs and overfitting of models [11].

Many feature reduction (FR) techniques have been developed to counter these issues. These techniques reduce the dimensionality of data by eliminating the features which are not correlated or useful in finding the outcome variable. Past SCP studies have applied filter-based ranking, wrapper-based subset selection, neural-network based, metaheuristic and extraction-based methods to find effective feature representations [2, 4, 12-16].

Since many diverse metrics have been used and a large variety of FR techniques are available to select the best among these metrics, it is crucial to analyse and compare them. This observation motivates us to conduct an extensive comparative study of 11 FR techniques in SCP. We use relevant Object-Oriented (OO) metrics and Graph-based metrics as independent features since they are the 2 most popularly used metric suites in SCP [2-4, 13, 17]. The experiments are conducted on 6 new datasets that we have collected through proprietary tools. AUC-ROC is used along with statistical tests for evaluation. The objective of our study is three-fold. Firstly, this study will help in the identification of the most important metrics in SCP. Secondly, this study will help in the identification of the most effective FR techniques for SCP. Finally, we can observe how the utility of these metrics and FR techniques varies with changes in the classification techniques.

Broadly, our study answers the following research questions (RQs) -

1. What is the significance of using FR techniques in SCP?
2. Which FR technique is most effective in SCP? Does the type of classifier influence the choice of FR technique?
3. Which features are most important in SCP?

The major contributions of this study are -

1. The study evaluates the performance of 11 FR techniques for SCP. These 11 techniques belong to

978-1-7281-8710-5/21/$31.00 ©2021 IEEE
DOI 10.1109/MSR52588.2021.00015

diverse categories of FR techniques like filter-based, wrapper-based, extraction-based, neural-network based, and metaheuristic techniques. To the best of the author's knowledge, this is the first extensive study that compares FR techniques in SCP.

2. The study is conducted on 6 new Java datasets that we have collected using open-source and proprietary tools. These datasets have OO metrics and graph-based metrics as independent features. The datasets are diverse in terms of domain, size, source of origin, and class-ratio. Further, the use of 3 classification techniques - Support Vector Machine (SVM), Naive Bayes (NB), and Decision Tree (DTree) with all combinations of 11 FR techniques and 6 datasets increases the generalizability of our results.
3. To establish the basic performance of the classifiers we use AUC-ROC as the performance measure since it is the recommended choice for the imbalanced SCP datasets. To increase the significance of our findings, the Friedman test along with the post-hoc double Scott-Knott test is used for further validation of performance of the FR techniques.

The rest of the paper is organized into 6 sections. Section II describes the past SCP studies and the various FR used by them. Section III provides details about the 11 FR techniques that are compared in this study. Section IV elaborates the experimental framework which includes the data collection procedure, independent features - OO metrics and graph-based metrics, classification techniques, performance indicators, statistical tests, and the overall experimental setup. Section V presents the experimental findings and answers the RQs mentioned above. Section VI presents the threats to validity. Section VII concludes the study and future work.

II. LITERATURE REVIEW

For doing software-related predictive modelling we first study the software project, process, and product, and then extract the metrics that describe them. Malhotra and Bansal [17] reviewed SCP studies and found that product-based metrics such as CK [10], Lorenz and Kidd [18], MOOD [19], QMOOD [20], etc. have been predominantly used in this field for model training. Malhotra et al. [2] surveyed 38 SCP studies and found that most of them used OO metrics for software change prediction. Catolino et al. [8] proposed a study with the use of developer-related metrics to enhance the prediction performance in SCP by 22% as compared to single-version models. Kagdi and Melatic [7] proposed a study that focuses on the need to combine single-version and evolutionary dependencies based on multiple versions. The study concluded that the combination improves the overall performance of SCP. Yang et al. [21] proposed a different approach for training SCP models by using metrics based on the graphical structure of the software. Bhattacharya et al. [6] reflected the utility of graph-based metrics for software predictive modelling fields like bug severity, maintenance, and defect prediction. Giger et al [22] advocated using a combination of OO and graph-based metrics.

To cover all the aspects of modern software, a large set of features are considered for model training. This increase in the feature set size causes problems related to high data dimensionality like unsatisfactory results of predictive classifiers [23, 24, 25]. These studies suggest the use of FR techniques to eliminate repetitive and irrelevant features.

Blum and Langley [26], Tang et al. [27], Chandrashekar, and Sahin [28] have given invaluable insights surveying and explaining the work in ML on FR techniques. FR techniques can be broadly classified into different groups based on the approach and framework followed by them. Some of them are -filter-based methods, wrapper-based methods, feature extraction-based methods, evolutionary methods, and neural network-based methods. Malhotra et al. [2] conducted a review of primary studies on SCP and found 58% of them were using FR techniques. Chi-Square test [8], CFS [9], wrapper method [14], PSO [15], Genetic Algorithm [16], PCA [23] are some of the techniques that have been previously used in SCP for the specified purpose.

Since a large variety of metrics and FR techniques have been applied in SCP, there is a need to compare them to identify the most effective ones. Many such comparative studies are already present for software related fields like software risk analysis, software cost estimation, key class identification, and software defect prediction. On the contrary, no comparative study providing definitive guidance on FR techniques is present for SCP.

On the lines of works by Shivaji et al. [24] and Xu et al. [25] on the comparison of FR techniques in SDP, we propose a similar study in SCP where we investigate the effects of 11 FR techniques using NB, SVM, and DTree classifiers over 6 Java projects - ant, hibernate, wro4j, tomcat, javaclient, and neuroph.

III. FEATURE REDUCTION TECHNIQUES

High-dimensionality datasets often create various issues like high computational cost, reduction of the impact of important features in the model, and low prediction performance due to overfitting [11]. FR techniques counter these problems by selecting, removing, or transforming the original features to reduce the dimensionality of the dataset [24, 25]. This study compares 11 FR techniques belonging to different categories. These techniques are listed in Table I. These techniques are explained in the subsections below.

TABLE I. FR TECHNIQUES

Category	FR Technique	Abbreviation
Filter-based	Chi-Square	CS
	Gini Index	Gini
	ReliefF	RF
	Fisher Score	FS
Wrapper-based	Forward Learning	Wrap
Extraction-based	Principal Component Analysis	PCA
	Linear Discriminant Analysis	LDA
Metaheuristic	Particle Swarm Optimization	PSO
	Artificial Bee Colonization	ABC
	Firefly Algorithm	FF
Neural-Network based	Auto-Encoders	AE

A. Filter-Based Techniques

Filter-based techniques reduce the number of features by providing a rank to the original features according to some statistical or probability-based measure. The features with ranks higher than the threshold are chosen while the others are eliminated in the reduced dataset. This threshold value is an external variable. In this study, we have used four different feature ranking mechanisms –

1) Chi-Square: CS [12] tests the dependency of features with the dependent variable and selects those with the highest dependence. Higher the CS value for a feature, the more effective it is in the prediction of the dependent variable.

2) Gini Index: Gini [29] is a splitting criterion that is used in DTrees. Gini calculates the entropy of the dependent variable with each feature. Gini is inversely related to the effectiveness of the feature.

3) ReliefF: RF [25] is a supervised multivariate feature selection (FS) technique that assigns a feature score to each feature. A feature is assigned a higher score if its value remains similar across samples of the same class and changes across samples of different classes. Features having higher scores are chosen for model training.

4) Fisher Score: FS [25] is a statistical method for calculating the rank of each feature. A Higher Fisher score implies a higher priority of a feature. FS is calculated as follows:

$$\text{FS} = \frac{n_p(\mu_p + \mu_T)^2 + n_N(\mu_N - \mu_T)^2}{{\sigma_T}^2} \quad (1)$$

where μ_T is the mean of the values in the dataset, σ_T is the standard deviation of the dataset, N signifies negative samples and p signifies positive samples for the original dataset.

B. Wrapper-Based Techniques

Wrapper-based techniques [14] use classification techniques to select the best feature subset by training for different combinations of features. Wrapper-based techniques use forward learning, backward elimination, or a combination of both for searching for the best combination of features. Wrapper-based techniques follow a greedy-search approach where the classification accuracy is the evaluation measure for every feature combination. In this study 3 classification techniques have been used for wrapper-based FS - NB, DTree, and SVM.

C. Extraction-Based Techniques

Extraction-based techniques project the high-dimensional original dataset into a lower-dimensional subspace. This projection is formed by extracting significant structures and relationships from the original features. These transformed features can be used for training the model. Two such techniques have been used in this study for comparison -

1) Principal Component Analysis: PCA [23] represents the dataset in the form of orthogonal vectors called principal components. These principal components are formed by linear combinations of the original features. For a dataset with n number of features, n principal components are created. A set of k principal components having the highest variance are selected as the new features. Here, k is the reduced dimension size.

2) Linear Discriminant Analysis: LDA [30] is a supervised technique that creates new features considering the dependent variable. The objective of LDA is to project the original dataset to a subspace where the distance between similar class samples is minimized while the distance between samples of different classes is maximized.

D. Metaheuristic Techniques

Metaheuristic techniques are based on the real-life evolution strategies adopted by various living organisms. Their use for feature reduction has recently increased due to their ability to find a globally optimum combination of features in less time. For all the described techniques, AUC-ROC was taken as the fitness function. In this study, 3 metaheuristic techniques have been used for comparison -

1) Particle Swarm Optimization: PSO [15] is a metaheuristic computation technique that has been widely applied in feature selection due to its computational efficiency, fast convergence and its ability to find global optimum solutions. PSO consists of a group of particles moving in the solution space of the problem. Every particle's position represents a possible feature subset. A fitness function evaluates the suitability of a solution for the given problem. The particles move to better positions in every iteration based on their personal experience and the overall experience of the population. The process terminates on achieving the desired fitness function value or after a fixed number of iterations.

2) Artificial Bee Colonization: ABC [31] is a swarm-based metaheuristic technique that models the food-search behavior of bee colonies. In ABC, potential solutions to the optimization problem are represented by position of food sources. In feature selection problem, these positions represent a possible feature subset. The fitness function (the nectar amount at a food source position) represents the AUC-ROC value achieved considering the feature subset of that solution. The employed, onlooker, and scout bees perform various iterations to reach the best food source position. The process terminates on achieving the desired fitness function value or after a fixed number of iterations.

3) Firefly Algorithm: FF [32] is a metaheuristic technique based on the flashing behavior of the fireflies. The position of fireflies represents potential solutions to the optimization problem. The intensity of the firefly is the value of the fitness function on that position. On each iteration, fireflies move to better position depending on the bioluminescent attraction between them. The process terminates on achieving the desired fitness function value or after a fixed number of iterations.

E. Neural Network-Based Techniques

AE [30] is an artificial neural network that encodes the original feature set into a compressed and efficient feature set. AE is an unsupervised learner with a bottleneck architecture

with input features as the original feature set and output features as the encoded feature set. AE aims to reduce noise, eliminate uncorrelated features, and construct the new feature set as close to the original feature set as possible.

IV. EXPERIMENTAL SETUP

The experimental framework used in the study is depicted in Fig. 1. All the individual components are explained in further subsections.

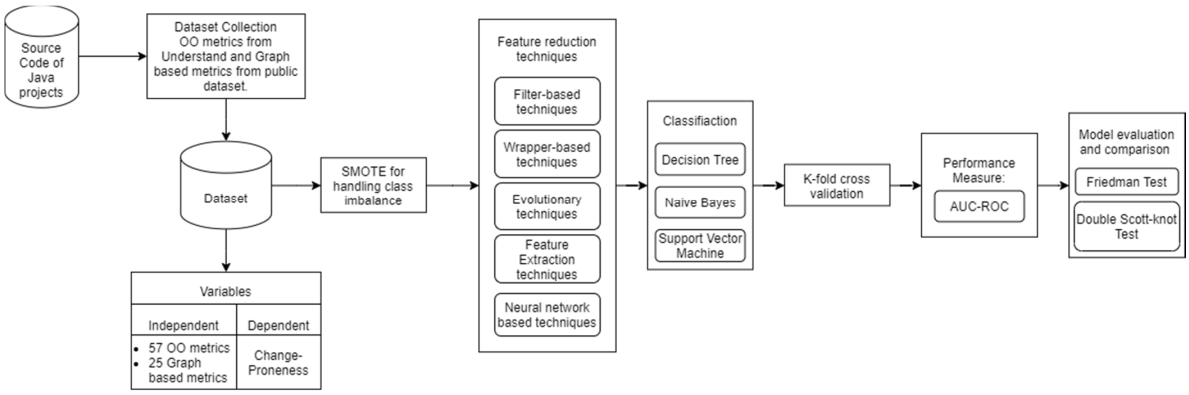


Fig. 1. Experimental Framework

A. Variables

1) Independent variables: In this study, two categories of independent variables are used - OO metrics and graph-based metrics. Graph-based metrics are further divided into core metrics and centrality measure metrics. OO metrics [10,18-20] describe the class structure like lines of code, number of methods, number of attributes, cyclomatic complexity, etc. They also represent the relationship between the classes like coupling, cohesion, etc. This study uses 60 handpicked OO metrics[1] collected using the UNDERSTAND tool (http://www.scitools.com/). The graph-based metrics are derived from a dependency graph where each class represents a node and the inter-class relationships like inheritance, containership, etc. form the edges [33]. Core graph metrics calculate features like in-degree, out-degree, weighted connections, etc. of a class, and centrality measures like PageRank, K-Coreness, Betweenness, etc. indicate the importance of a class in the software. This study uses 26 graph-based metrics taken from a publicly available dataset [33]. A total of 86 independent variables have been used in this study for classification. These 86 variables are listed in Table II.

2) Dependent Variables: Change-proneness of a class is the binary dependent variable in this study. There are various types of changes in a class like addition or removal of methods and attributes, modifications in the class declaration, etc. In this study, we model SCP as a binary-classification problem hence all these changes are labeled in the same category i.e. change (1). If no modification has been done in a class between the two subsequent releases, then the sample is labeled as no change (0).

TABLE II. METRICS COMPRISING DATASETS

Category	Metrics included in the study
OO metrics	Average Cyclomatic Complexity (F1), Average Modified Cyclomatic Complexity (F2), Average Strict Cyclomatic Complexity (F3), Average Essential Cyclomatic Complexity (F4), Average Number of Lines (F5), Average Number of Blank Lines (F6), Average Number of Lines of Code (F7), Average Number of Lines with Comments (F8), A Base Classes (F9), Coupling Between Objects (F10), Count of Modified Coupling Between Classes (F11), Number of Children(F12), Classes (F13), Class Methods (F14), Class Variables (F15), Executable Unit (F16), Number of Files (F17), Function (F18), Instance Methods (F19), Instance Variables (F20), Local Methods F(21), CountDeclMethodAll (F22), Local Default Visibility Methods F(23), Private Methods (F24), Protected Methods (F25), Public Methods (F26), Inputs (F27), Physical Lines (F28), Blank Lines of Code (F29), Lines of Code (F30), Declarative Lines of Code (F31), Executable Lines of Code (F32), Lines with Comments (F33), Outputs (F34), Paths (F35), Paths Log(x) (F36), Semicolons (F37), Statements (F38), Declarative Statements (F39), Executable Statements (F40), Cyclomatic Complexity (F41), Modified Cyclomatic Complexity (F42), Strict Cyclomatic Complexity (F43), Essential Complexity (F44), Knots (F45), Max Cyclomatic Complexity (F46), Max Modified Cyclomatic Complexity (F47), Max Strict Cyclomatic Complexity (F48), Max Essential Complexity (F49), Max Knots (F50), Depth of Inheritance Tree (F51), Nesting (F52), Minimum Knots (F53), Lack of Cohesion in Methods (F54), Modified Lack of Cohesion in Methods (F55), Comment to Code Ratio (F56), Sum Cyclomatic Complexity (F57), Sum Modified Cyclomatic Complexity (F58), Sum Strict Cyclomatic Complexity (F59), Sum Essential Complexity (F60)
Graph-based metrics	Weighted Incoming Dependencies (F61), Weighted Outgoing Dependencies (F62), Total Weighted Dependencies (F63), PageRank on Directed Graph (F64), PageRank on Unweighted Graph (F65), PageRank on Undirected, Weighted Graph (F66), PageRank on Weighted Graph With Back Recommendations (F67), PageRank on Weighted Graph (F68), Authority Value on Directed Unweighted Graph (F69), Authority Value on Directed, Weighted Graph (F70), Hub Value on Directed Unweighted Graph (F71), Hub Value on Directed Unweighted Graph (f72), Betweenness Value on Undirected Unweighted Graph (F73), Betweenness Value on Undirected Weighted Graph (F74), Core Number on

	Undirected, Unweighted Graph (F75), Core Number on Weighted Graph (F76), Top Classes That Are Incoming Dependencies (F77), Top Classes That Are Outgoing Dependencies (F78), Total Top Class That Are Dependencies (F79), Total Weight of Incoming Dependencies From Top Classes (F80), Total Weight of Outgoing Dependencies From Top Classes (F81), Total Weight of All Dependencies From Top Classes (F82), Betweenness Value on Weighted Graph With Back Recommendation (F83), Incoming Dependencies (F84), Outgoing Dependencies F(85), Total Dependencies (F86).

B. Dataset

This section describes the procedure of dataset collection and the properties of the collected datasets. 6 open-source Java projects are chosen for collecting the datasets [34]. For each project, the source code of two subsequent releases is collected from GitHub (https://github.com) or Sourceforge (https://sourceforge.net). From the source code of both releases, 60 OO metrics[1] for JAVA classes are collected using the UNDERSTAND tool (http://www.scitools.com/). Now the two releases of the project are compared to identify the classes in which there have been some changed or modified classes. Firstly, the classes from both releases are compiled together in the dataset. The classes which are not common to both the releases were classified as changed. For the remaining common classes, the number of lines increased, decreased, and modified from first to the second release are found. For every class, if no lines were added, deleted, or modified it is labeled as no change (0) else change-prone (1). Following the above steps, we get our data with OO metrics and the dependent variable. Now the 26 graph-based metrics available as publicly accessible datasets [33] are combined with the OO metric data collected in previous steps though inner join over the columns containing class names in both datasets This completes our data collection procedure. Finally, we have 6 datasets with 86 independent variables and 1 dependent variable described in Section IV-A. These datasets and their statistics are shown in Table III.

TABLE III. STATISTICS OF DATASETS

Dataset	Release 1	Release 2	Number of classes	% of changed classes	Description
Ant	1.6.0	1.6.1	319	10.3	Java-based build tool of Apache
Tomcat	9.0.3	9.0.4	354	1.97	Open-source implementation of Java Servlet by Apache
Wro4j	1.6.2	1.6.3	169	24.7	Java project that combines web tools
Hibernate	5.2.11	5.2.12	2022	6.77	Object Relational Mapping (ORM) tool
Javaclient	2.2.0.0	2.2.0.1	211	4.24	Appium project on Java
Neuroph	2.1.1	2.2	112	26.5	Java class library and GUI tool

C. Classifiers

In this study, 3 classification techniques have been used for the prediction of change-prone classes. These classifiers have been widely used in the previous comparative studies [24, 25]. Also, these techniques have proven to have good prediction performance and are stable for imbalanced data [35]. These techniques are very diverse from each other in terms of learning strategies and have fast computation efficiency [36]. These classifiers are explained below-

1) Naive Bayes - NB [36] is a statistical probability-based ML technique that predicts the probability of occurrence of a sample in different classes using prior and posterior probabilities and then labels the sample with the class having maximum probability.

2) Support Vector Machine - SVM [36] is a popular supervised ML technique often used in classification and regression problems. SVM develops a hyperplane to divide the sample space such that the distance of the hyperplane with the closest sample is the largest.

3) Decision Tree - DTree [36] constructs a tree where each node represents a condition on a feature and each path from the root to leaf node leads to an output label. While moving from root to leaf various decisions take place according to the features of the samples and the conditions on those features. This is a widely used technique in various problem domains.

D. Performance Measures

In this study, we use AUC-ROC for evaluating the performance of the different techniques. AUC-ROC has been widely used in past studies on SCP [3, 5, 13, 14]. AUC-ROC measures the capability of a model to distinguish between change-prone and no-change classes. It is a summary of the ROC curve which is a plot between True Positive Rate (proportion of change-prone classes that were identified correctly as change-prone) and False Positive Rate (proportion of no-change classes that were identified as change-prone). AUC-ROC values range between 0 to 1 where 1 means that the model is identifying every sample correctly and 0 means that the model is performing poorly and reciprocating the actual class labels. AUC-ROC is a highly effective performance measure especially when the dataset is imbalanced, which is often the cases with datasets in SCP.

E. Statistical Tests

This study uses the non-parametric Friedman test [37] followed by the post-hoc double Scott-Knott test [38] for a consolidated performance comparison of different FR techniques across datasets. The tests are conducted at the significance level (α) of 0.05. To conduct these tests the following steps are followed -

1) Formulation of the null hypothesis (H_0) - H_0 states that all the FR techniques are similar in performance in SCP. In other words, the use of one FR technique over another does not affect the classification performance.

2) Computing Average Friedman Ranking (AFR) - Consider we have f FR techniques and d datasets. For every FR technique, the average rank over d datasets is conducted based on performance measures described in Section IV-D. The formula for calculating AFR of i-th FR technique is –

$$AFR_i = \frac{1}{d}\sum_{j=1}^{d} R_{ij} \quad (2)$$

where R_{ij} is the rank of the i-th technique for the j-th dataset.

3) Computing Friedman test statistic (FTS) - Using the AFR of the FR techniques computed in Step (2), FTS is computed as –

$$FTS = \frac{12d}{f(f+1)}\left(\sum_{i=1}^{f} AFR_i^2 - \frac{f(f+1)^2}{4}\right) \quad (3)$$

4) Accept or Reject H_0 - If the FTS computed in equation (3) is greater than the chi-square distribution value for significance level (α) and f-1 degree of freedom, the null hypothesis H_0 is rejected else H_0 is accepted. if H_0 is rejected, we conduct the post-hoc test according to step (5)

5) Scott-Knott test - It creates distinct clusters of FR techniques using hierarchical clustering. Clusters are created based on a performance metric of FR techniques for a given dataset. Ranks are assigned to techniques based on their cluster.

6) Double Scott-Knott test - It is used for comparison of the techniques across datasets. For each technique, we obtain 6 Ranks using the Scott-Knott test each corresponding to a unique dataset. Now, using dataset wise ranks as an input, we conducted a second Scott-Knott test to obtain overall ranks for an FR technique.

F. Experimental Procedure

The experiments are performed using the datasets described in Table III. Since the datasets are highly-imbalanced, we have used Synthetic-Minority Oversampling (SMOTE) [39] to counter this issue. We used a SMOTE ratio of 0.6 [40].

Next, we have applied the FR techniques given in Table I using the parameters described here. Let p be the number of original features. The filter-based and wrapper-based methods are implemented using the Skfeature library of python. The number of reduced features is a variable in the FR techniques that depends on various factors like dataset distribution, the mathematical formula used by filter methods, etc. Hence, for each technique, we used the brute force technique for variable selection to optimize the percentage of feature reduction. We conducted experiments with fraction values in the range of 0.1 to 0.9 with a difference of 0.1 [41]. From these experiments we found out that 0.5*p is the optimal number of features to be taken for filter-based and wrapper techniques. For wrapper-based techniques, we have used a forward learning search strategy.

PCA and LDA are implemented using the Sci-kit learn python library. The number of components in PCA is chosen as 40 as we want 95% variance to be retained in the new features [42].

The parameters for PSO, ABC, FF, and AE are selected by running the algorithms multiple times on different parameter combinations. The metaheuristic techniques and AE are implemented from scratch in python. PSO is implemented with an initial population size of 30, maximum number of iterations to be 50, c1 and c2 with values 1.496 respectively, and the velocity range to be [-2,2]. ABC is implemented with an initial population size of 30, abandonment limit 100, upper/lower limit as [-1,1], and maximum iterations as 50. FF is implemented using population size 20, α=0.8, β=1, γ=0.7, and maximum iterations as 40.

For AE we use Adadelta optimizer, categorical cross-entropy loss function, and relu activation function. The number of epochs is 50 with a batch size of 100.

The classification techniques used in this study are mentioned in Section IV-C. They are implemented with default parameter settings of Sci-kit learn library, and a 10-fold cross-validation strategy is used for training and evaluation of performance. The performance is measured using the indicators described in Section IV-D. Finally, statistical tests are conducted which are explained in Section IV-E. This entire framework is depicted in Fig. 1.

V. EXPERIMENTAL RESULTS

A. RQ1 - What is the significance of using feature reduction techniques in SCP?

The objective of this RQ is to find the need to apply FR techniques in SCP. Here 11 sets of experiments are conducted considering each of the 11 FR techniques. In addition, a set of experiments is conducted without using FR techniques. Each set of experiments consist of 18 different combinations formed by the 6 datasets and 3 classifiers.

The AUC-ROC results obtained from these experiments are shown as box-plots in Fig. 2, 3, 4. The three box-plots are differentiated based on the three different classifiers (DTree, NB, and SVM) used in the experiments. Each of these box-plots shows the variation of AUC-ROC over different datasets for the 11 FR techniques and the one labelled 'No FR' corresponding to the case where the no FR technique is applied. In the box-plot for the SVM classifier, it is observed that all the FR techniques except the filter-based CS have a higher median AUC-ROC value than the case where no FR technique is used. A similar scenario is observed for the DTree classifier where the scenario of not using any FR gives the second-worst median AUC-ROC after AE. The results of the NB classifier are also against the scenario where we do not use any FR technique as here it gives the third-worst median. To summarize, across all the three classifiers the median AUC-ROC score of the best FR technique is higher than the median of no-FR case. Hence, by analysis of the AUC-ROC performance results, we can say that the application of FR techniques is necessary since there are a lot of redundant and unnecessary features that are not effective in SCP.

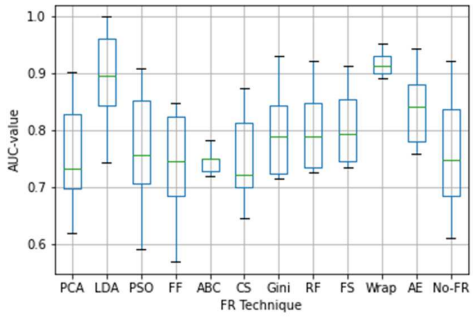


Fig. 2. Box plot of AUC-values over datasets using SVM classifier

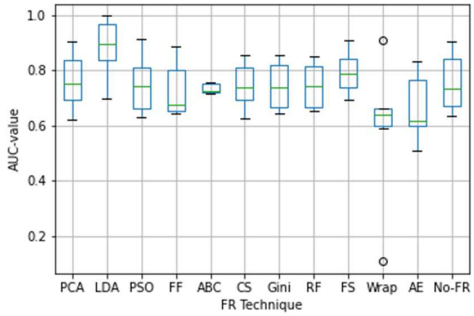


Fig. 3. Box plot of AUC-values over datasets using NB classifier

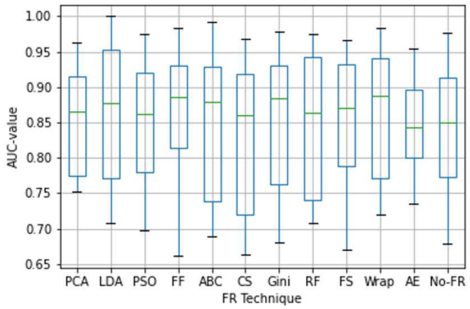


Fig. 4. Box plot of AUC-values over datasets using DTree Classifier

To further validate the inferences drawn from AUC-ROC results, we perform statistical validation. We first applied the Freidman tests where we got the Friedman Statistic Value as 12.295, 22.483, and 25.982 for DTree, NB, and SVM respectively. These values suggest that the initial hypothesis H_0 is rejected for all the three classifiers. This concludes that all 11 FR compared in the study have significant differences in prediction performance. Since we also compared a case where no FR technique was applied, it can be concluded from the rejection of H0 that there is a difference in performance between the scenario where FR techniques are used and where FR techniques are not used. The results of double Scott-Knott tests further justify these observations. The results of the double Scott-Knott test are displayed in figures 5, 6, and 7 using graphs. Here each graph is created based on a classifier used. For each graph, the x-axis hoists the status of the FR technique and the y-axis hoists average rank values. The height of each bar demarcates the average rank of each technique over six datasets. The colour of each bar indicates the distinct cluster to which the corresponding FR technique is assigned. For DTree, the scenario of not using any FR technique is ranked the lowest with an average rank of 8.5. For NB and SVM, the scenario of not using any FR technique is not the worst since there are a few FR techniques like PSO, CS, and FF that perform worse. Table IV summarizes the results of the double Scott-Knott. For each classifier, the cluster division is shown along with the techniques in that cluster and other cluster statistics like mean, median, and standard deviation. It is observed from Table IV that the scenario of not using any FR technique always falls in the cluster of the worst techniques which have the lowest rankings. Since the results of double Scott-Knott tests are independent of the original AUC-ROC values of a technique on the different projects, we can say that the scenario of not using any FR technique is overall undesirable.

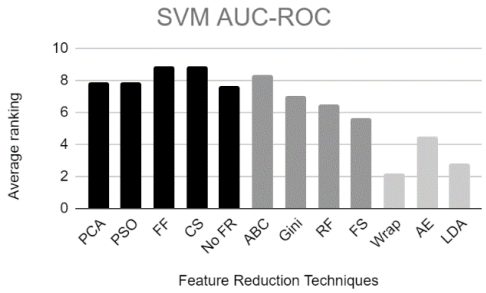


Fig. 5. Double Scott-knot test and Friedman ranking over datasets for SVM classifier

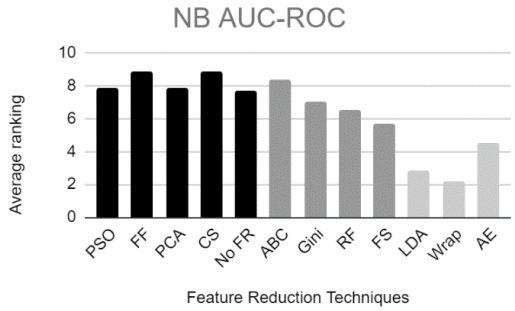


Fig. 6. Double Scott-knot test and Friedman ranking over datasets for NB classifier

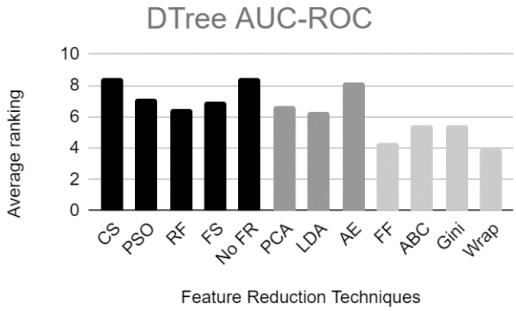


Fig. 7. Double Scott-knot test and Friedman ranking over datasets for DTree Classifier

Conclusion - The experimental results indicate the presence of irrelevant and redundant features that affect the prediction performance of the classifiers. The AUC-ROC values and the statistical tests strongly suggest the necessity of using FR techniques in SCP to select effective features for training predictive models.

TABLE IV. SUMMARY OF DOUBLE SCOTT-KNOTT TEST

	Performance Measure	Techniques	Average Rank	Median Rank	Std. Deviation
DTree					
1	AUC-ROC	FF, ABC, Gini, Wrap	4.8333	4.9165	0.7818
2		PCA, LDA, AE	7.0556	6.6670	0.9767
3		CS,PSO, RF, FS, No FR	7.5333	7.1667	0.9159
NB					
1	AUC-ROC	AE, LDA, Wrap	3.1667	2.8333	1.2018
2		ABC, Gini, RF, FS	6.8749	6.7500	1.1168
3		No FR,CS, PSO, PCA, FF	8.2000	7.8333	0.5821
SVM					
1	AUC-ROC	Wrap, AE, LDA	3.1666	2.8330	1.2019
2		ABC, Gini, RF, FS	6.8758	6.7500	1.1157
3		PSO, FF, CS, PCA, No FR	8.1998	7.8330	0.5820

B. *RQ2- Which feature reduction technique is most effective in SCP? Does the type of classifier influence the choice of feature reduction technique?*

In this RQ, we compare the performance of selected FR techniques in SCP. Also, we study the effect of different classifiers on the performance of the FR techniques.

For SVM, it is observed from Fig. 2 that the least AUC value achieved by Wrap is better than most of the other FR techniques except LDA. The wrapper technique shows good performance as AUC-ROC values over all the datasets lies within the narrow range of 0.89 to 0.93. This shows the superiority of Wrap over other FR techniques. ABC has a lower range of values for these datasets and has the lowest maximum as compared to other techniques. PSO and FF have shown a comparable performance with the highest median for PSO. This observation showed an efficiency of PSO over ABC and FF. All the four filter-based techniques have achieved almost equivalent AUC-ROC range with FS being the most efficient and CS being the least. AE has also performed well due to its mechanism of minimizing errors between input and output features. LDA has shown significantly better results than PCA. This observation indicates that the supervised nature of LDA makes it more efficient than the unsupervised technique PCA.

For NB, it is observed from Fig. 3 that the same trend between LDA and PCA has been followed which shows that this might be true that LDA performs better than PCA independent of the classification technique used in SCP. LDA has performed the best among all the FR techniques in terms of the median of AUC values. For the metaheuristic techniques, it is seen that for PSO and ABC, the median AUC-ROC lies above 0.7 which shows a great predictive efficiency. FF in turn has a comparatively worse median showing the superior performance of PSO and ABC in this case. Filter-based techniques especially CS, Gini, RF have shown equivalent median values. FS has shown the best performance among all the filter-based techniques. Wrap has extremes at very low and high AUC-ROC which shows a change in performance efficiency with datasets.

For DTree, it is observed from Fig. 4 that the minimum AUC value lies above 0.65 which shows that DTree is an effective classification technique. 3 out of 4 filter-based techniques i.e. CS, RF, and FS have achieved almost similar median AUC-ROC but the range of all of these techniques differ. This shows that some filter-based FR techniques are more efficient in selecting better features than the others. AE and FF have comparatively shown less range of values than others. LDA has given the best highest AUC-value as compared to any other techniques. In this case, the performance of all the FR techniques have been comparable unlike NB and SVM. This shows that classifiers widely affect the performance of FR techniques.

The results in Fig. 5, 6 and 7 conclude that some feature techniques are more effective and some are least effective irrespective of the classification technique used. For all three classifiers in Fig. 5, 6, and 7, PSO and No FR have been among the highest ranks, i.e. lowest prediction performance. For NB and SVM, LDA has been among the top performers with rank 2.83, which shows the capability of LDA in efficient feature reduction in SCP. But for DTree, LDA has achieved a rank of 6.33 which lies in the cluster with the second-highest average ranking showing a difference in its performance according to the classifier.

Table IV has shown the clustering of different FR techniques according to the double Scott-Knott test. One important observation is that at least two metaheuristic techniques always lie in the same cluster i.e. FF and ABC or PSO and FF. This shows that metaheuristic techniques tend to perform very similar to each other with less or no significant differences. However, these techniques have always been in the last or second last cluster showing the lower efficiency of metaheuristic techniques in SCP. Wrap technique always lies in the topmost cluster and has achieved the best rank in all the three classifiers. Wrap uses the classification efficiency of the classifier as the evaluation measure to extract the most efficient feature subset. This mechanism helps it in synergizing with corresponding classifiers. The filter-based techniques are divided into various clusters which show that each technique differs from the other in terms of FS.

Conclusion: There are some FR techniques that perform better than other FR techniques. LDA and Wrap have performed very efficiently while PSO and CS have performed worse overall. Filter-based techniques have shown similar results with each other. Metaheuristic techniques have shown good AUC-values but have achieved comparatively lower ranks. It is observed that some techniques have shown consistent results irrespective of the classifier used while the average ranking of various techniques like AE, ABC, FF, etc. differ according to the classifier which shows that classifiers affect the choice of FR techniques.

C. *RQ3- Which features are most important in SCP?*

In this competitive study, we have conducted a total of 204 different experiments to answer questions studying FR techniques in varied aspects. While doing these experiments we generated 48 unique feature rankings of 86 original features. These 48 unique feature rankings were obtained

using the 8 FR techniques (CS, Gini, RF, FS, Wrap, PSO, ABC, FF) over 6 unique datasets. AE, LDA, and PCA were not considered because they return new features extracted from the given feature set. Also, the ranks given to features by different FS techniques do not change with a change in classifiers so we have ignored that dimension while answering RQ3.

Now, on the basis of these 48 feature ranks for each feature, we have conducted a Scott -Knott test to form three priority clusters of features. Table V describes the clusters by showing the priority of each cluster and the feature codes of each feature in a cluster. The feature name to feature code mapping is stated in Table II. From studying Scott-Knott test results we have amassed interesting inferences regarding the importance of features.

TABLE V. RANKING OF METRICS USING SCOTT-KNOTT TEST

Cluster No.	Metric Importance	Metrics included in the cluster
1	High	F10, F12, F13, F14, F15, F18, F23, F24, F25, F26, F31, F46, F47, F48, F51, F52, F54, F55, F69, F71, F72, F75, F76, F77, F78, F79, F83, F84, F85, F86
2	Medium	F1, F2, F3, F4, F5, F9, F19, F20, F21, F27, F30, F33, F37, F38, F39, F40, F49, F50, F53, F56, F67, F68, F70
3	Low	F6, F7, F8, F11, F16, F17, F22, F28, F29, F32, F34, F35, F36, F41, F42, F43, F44, F45, F57, F58, F59, F60, F61, F62, F63, F64, F65, F66, F73, F74, F80, F81, F82

Cluster 1 has the highest priority features and it consists of 30 features, 30% of the OO metrics and 46.15% of the graph-based metrics belong to this cluster indicating the higher importance of graph-based metrics. The Maximum complexity value-based metrics (F46, F47, F48), cohesion and coupling based metrics (F10, F54), class access type based metrics (F24, F25, F26), basic dependency-based metrics (F84, F85, F86), hub-value related metrics (71, 72) and core-value related metrics(75, 76) are some of the broad categories which predominantly constitute cluster 1.

Cluster 2 has medium priority and it consists of 23 features, 33.33% of the OO metrics and 11.54% of the graph-based metrics belong to this cluster. It is hard to identify proper groups of metrics belonging to this cluster as it lies between the cluster with the highest importance and lowest importance. However, Averages of different complexities in OO metrics (F1, F2, F3, F4), some metrics related to Lines of Code (F5, F30, F33, F56), metrics related to statements (F37, F38, F39, F40), and some of the page-rank based metrics (F67, F68) are identifiable metrics groups in this cluster.

Cluster 3 has the lowest priority and consists of 33 features 38.33% of the OO metrics and 42.31% of the graph-based metrics belong to this cluster. The metrics belonging to this cluster are seldom picked by FS techniques for training classifiers. Sum of different complexities in OO metrics (F1, F2, F3, F4, F7), instance variable-based metrics (F15, F16, F17, F19), some of the metrics based on a weighted graph representation of the software (F61, F62, F63, F64), and some of the top class-related metrics (F80, F81, F82) are a part of this cluster.

Conclusion: From Table V we can identify the features and feature groups that have more importance in SCP. 38 of the 60 OO metrics and 15 of the 26 graph-based metrics belong to the first two clusters and are frequently shortlisted by FS techniques. However, the other metrics making up cluster 3 are not necessarily irrelevant; it might also mean that the information they are conveying is also provided by other metrics.

VI. THREATS TO VALIDITY

A. Internal Validity - Internal validity deals with the bias in the various choices we make in our study that can affect the validity of the cause-and-effect relationships drawn from it. In this study, we used OO metrics and graph-based metrics as independent variables for predicting change-proneness of classes. These metrics are very popular and effective in past SCP studies. The 11 FR techniques compared in this study are taken from the various categories to make the study more generalized and extensive. The choice of NB, DTree and SVM was based on their popularity, effectiveness and diversity of prediction algorithms utilized by them. However, this can be a source of bias and hence more classifiers from categories like neural-networks, ensembles should be used.

B. Construct Validity - Threats to construct validity are concerned with bias in the experimental design used in the study. To eliminate this bias, we have described the complete experimental setup in Section IV. We used SMOTE which is a widely recommended technique to deal with class imbalance. We have mentioned the implementation details of every FR technique used in this comparative study. The parameter settings used for each of these FR techniques are carefully chosen through extensive experiments and suggestions from previous studies. However, a more careful experiment specifically for choosing those hyper-parameters where no guidelines are mentioned. We used AUC-ROC as a performance measure due to its wide usage in similar studies especially for imbalanced datasets. Nonetheless, other performance measures like F-Measure can also be employed here.

C. Conclusion Validity - Conclusion Validity deals with the reliability of the experimental observations and results drawn from the study. We have used the popular Freidman test followed by the double Scott-Knott test at a significance level of 0.05 to evaluate the performance of the FR techniques. The use of 6 diverse datasets and 3 classification techniques makes this study extensive and ensures that all the conclusions drawn in the study are significant and not just trivial effects.

D. External Validity - External Validity deals with the generalization of our experimental findings. In this study, we generated 6 new datasets derived from open-source Java projects. We collected this dataset from two sources, one is the popular proprietary tool UNDERSTAND and other is an open-source research dataset. Although the datasets are

diverse in terms of the domain, code structure, size etc., we cannot claim that our findings work with datasets with a different set of metrics or if they are project-level or module-level.

VII. Conclusion And Future Work

In this study, we compared the performance of 11 FR techniques in SCP using OO and graph-based metrics. Extensive experiments were conducted on 6 newly generated Java datasets and across 3 classifiers. Using AUC-ROC performance measure along with the statistical tests for evaluation and comparison, the major conclusions drawn from the experimental findings are –

1) FR techniques are effective in SCP. They help in the extraction of features that are useful in predicting change-prone classes. The use of such features increases the predictive performance of classifiers.

2) There are significant differences in the performance of the 11 FR techniques compared in the study. While LDA, AE, and Wrapper-based FR techniques achieve better performance across different datasets and classifiers., FR techniques like PSO, PCA, and some filter-based techniques have shown unsatisfactory results.

3) The type of classifier used in SCP also affects the choice of FR techniques. The performance rankings of FR techniques like ABC and FF vary considerably with a variation in the classifier they are used with.

4) Some features have a higher probability of being chosen by FR techniques over others for model training in SCP. In the OO metric suite, coupling, cohesion, cyclomatic complexity, and lines of code are highly preferred. In graph-based metrics, inter-class dependencies and centrality measures like K-Coreness are highly effective.

In the future, we will extend this study by including more categories of FR techniques like hybrid and clustering-based techniques. The inclusion of more metrics like change-history of class and developer-related factors is another possible extension.

VIII. References

[1] K.K. Aggarwal and Y. Singh, “Software engineering”, New Age International (P) Limited, 2008.

[2] R. Malhotra and M. Khanna, “Software Change Prediction: A Systematic Review and Future Guidelines", In e-Informatica Software Engineering Journal, vol. 13, no. 1, pp. 227–259, 2019. doi: 10.5277/e-Inf190107.

[3] R. Malhotra and M. Khanna, “Mining the impact of object oriented metrics for change prediction using Machine Learning and Search-based techniques”, International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2015. doi:10.1109/icacci.2015.7275614

[4] R. Malhotra and M. Khanna, “An exploratory study for software change prediction in object-oriented systems using hybridized techniques”, Automated Software Engineering, vol. 24, issue 3, pp. 673–717, 2016. doi:10.1007/s10515-016-0203-0

[5] R. Malhotra, R. and A.J. Bansal, “Software change prediction: a literature review”, International Journal of Computer Applications in Technology, vol. 54, no. 4, pp. 240-256, 2016. doi:10.1504/ijcat.2016.080487

[6] P. Bhattacharya, M. Iliofotou, I. Neamtiu and M. Faloutsos, "Graph-based analysis and prediction for software evolution," 2012 34th International Conference on Software Engineering (ICSE), pp. 419-429, 2012. doi: 10.1109/ICSE.2012.6227173.

[7] H. Kagdi, and J.I. Maletic, “ Combining Single-Version and Evolutionary Dependencies for Software-Change Prediction”, Fourth International Workshop on Mining Software Repositories, 2007. doi:10.1109/msr.2007.2

[8] G. Catolino, F. Palomba, A. De Lucia, F. Ferrucci, and A. Zaidman, “Enhancing change prediction models using developer-related factors”, Journal of Systems and Software, 143, pp. 14–28, 2018. doi:10.1016/j.jss.2018.05.003

[9] M. Jureczko, “Significance of Different Software Metrics in Defect Prediction”, Software engineering : an international Journal (SeiJ), Vol. 1, no. 1, 2011.

[10] S. Chidamber and C. Kemerer, “A metric suite for object oriented design,” IEEE Transactions on Software Engineering, vol. 20, no. 6, pp. 476-493, 1994.

[11] Z. Xu, J. Liu, Z. Yang, G. An and X. Jia, “The Impact of Feature Selection on Defect Prediction Performance: An Empirical Comparison”, IEEE 27th International Symposium on Software Reliability Engineering (ISSRE), 2016. doi:10.1109/issre.2016.13

[12] L. Kumar, S. Lal, A. Goyal, and N. Murthy, “Change-proneness of object-oriented software using combination of feature selection techniques and ensemble learning techniques,” in Proceedings of the 12th Innovations on Software Engineering Conference. ACM, pp. 1-11, 2019. doi: 10.1145/3299771.3299778

[13] R. Malhotra and M. Khanna, “Investigation of relationship between object-oriented metrics and change proneness,” International Journal of Machine Learning and Cybernetics, vol. 4, no. 4, pp. 273–286, 2013. doi: 10.1007/s13042-012-0095-7

[14] G. Catolino and F. Ferrucci, “An extensive evaluation of ensemble techniques for software change prediction,” Journal of Software: Evolution and Process, vol. 31, issue 9, 2019. doi: 10.1002/smr.2156

[15] R. Malhotra and M. Khanna, “Software change prediction using voting particle swarm optimization based ensemble classifier”, in Proceedings of the Genetic and Evolutionary Computation Conference Companion, ACM, pp. 311-312, 2017. doi: 10.1145/3067695.3076007

[16] L. Kumar, R.K. Behera, S. Rath, and A. Sureka, “Transfer learning for cross-project change-proneness prediction in object-oriented software systems: A feasibility analysis,” ACM SIGSOFT Software Engineering Notes, vol. 42, no. 3, pp. 1–11, 2017. doi: 10.1145/3127360.3127368

[17] R. Malhotra and A. Bansal, "Predicting change using software metrics: A review", 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions), pp. 1-6, 2015. doi: 10.1109/ICRITO.2015.7359253.

[18] M. Lorenz and J. Kidd, “Object-oriented software metrics: A practical guide,” Prentice-Hall, Inc., 1994.

[19] F. B. Abreu and R. Carapuça, "Object-oriented software engineering: Measuring and controlling the development process," in proceedings of the 4th International Conference on Software Quality, 1994.

[20] J. Bansiya and C. G. Davis, "A hierarchical model for object-oriented design quality as-sessment," IEEE Transactions on Software Engineering, vol. 28, pp. 4-17, 2002.

[21] Y. Yang, J. Ai and F. Wang, "Defect Prediction Based on the Characteristics of Multilayer Structure of Software Network," IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C), pp. 27-34, 2018. doi: 10.1109/QRS-C.2018.00019

[22] E. Giger, M. Pinzger and H.C. Gall, (2012), “Can we predict types of code changes? An empirical analysis”, 9th IEEE Working Conference on Mining Software Repositories (MSR), 2012. doi:10.1109/msr.2012.6224284

[23] T. Quah and M. Thwin, “Application of neural networks for software quality prediction using object-oriented metrics”, International Conference on Software Maintenance (ICSM), 2003. doi:10.1109/icsm.2003.1235412

[24] S. Shivaji, S, E.J. Whitehead, R. Akella and S. Kim, “Reducing Features to Improve Code Change-Based Bug Prediction”, IEEE Transactions on Software Engineering, vol. 39, no. 4, pp. 552–569, 2012.

[25] Z. Xu, J. Liu, Z. Yang, G. An and X. Jia, “The Impact of Feature Selection on Defect Prediction Performance: An Empirical

Comparison", IEEE 27th International Symposium on Software Reliability Engineering (ISSRE), 2016. doi:10.1109/issre.2016.13

[26] A. L. Blum and P. Langley, "Selection of relevant features and examples in machine learning", Artificial Intelligence, vol. 97, issues 1–2, pp. 245-271, 1997. doi: 10.1016/S0004-3702(97)00063-5

[27] J. Tang, S. Alelyani and H. Liu, "Feature Selection for Classification: A Review" in Data Classification: Algorithms and Applications, CRC Press, 2014.

[28] G. Chandrashekar and F. Sahin, "A survey on feature selection methods", Computers & Electrical Engineering, vol. 40, issue 1,pp. 16-28, 2014. doi : 10.1016/j.compeleceng.2013.11.024.

[29] W. Duch, "Filter Methods", In: Guyon I., Nikravesh M., Gunn S., Zadeh L.A. (eds) Feature Extraction. Studies in Fuzziness and Soft Computing, vol. 207, 2006. doi: 10.1007/978-3-540-35488-8_4

[30] R. Malhotra and K. Khan, "A Study on Software Defect Prediction using Feature Extraction Techniques", 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2020. doi:10.1109/icrito48877.2020.9197999

[31] M. Schiezaro and H. Pedrini, "Data feature selection based on Artificial Bee Colony algorithm", J Image Video Proc, vol. 47, 2013. doi: 10.1186/1687-5281-2013-47

[32] M. Anbu and G.S. Anandha Mala, "Feature selection using firefly algorithm in software defect prediction", Cluster Computing, pp. 10925–10934, 2017. doi:10.1007/s10586-017-1235-3

[33] I. Sora and C. Chirila, "Finding key classes in object-oriented software systems by techniques based on static analysis", Information and Software Technology, vol. 116, 2019. doi: 10.1016/j.infsof.2019.106176

[34] [dataset] https://figshare.com/s/df5529168b9641bdd96e

[35] A. Balogun, S. Basri, S. J. Abdulkadir, V.Adeyemo, "Software Defect Prediction: Analysis of Class Imbalance and Performance Stability", Journal of Engineering Science and Technology 14(6), pp 3294-3308, 2019.

[36] E. Alpaydin, Introduction to Machine Learning, Cambridge, MA, USA: MIT press, 2014.

[37] M. Friedman, "A comparison of alternative tests of significance for the problem of m rankings", The Annals of Mathematical Statistics, vol. 11, pp. 86–92, 1940. doi: 10.1214/aoms/117773194

[38] B. Ghotra, S. McIntosh, and A. E. Hassan, "Revisiting the impact of classification techniques on the performance of defect prediction models", In Proceedings of the 37th International Conference on Software Engineering (ICSE), pp. 789-800, 2015. doi: 10.1109/ICSE.2015.91

[39] N. V. Chawla , K. W. Bowyer , L. O. Hall and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique", Journal of Artificial Intelligence Research, vol.16, pp. 321-357, 2002. doi: 10.1613/jair.953

[40] M. Lv, Y. Ren, and Y.Chen, "Research on imbalanced data based on SMOTE-AdaBoost algorithm", 3rd International Conference on Electronic Information Technology and Computer Engineering (EITCE), 2019. doi: 10.1109/EITCE47263.2019.9094859

[41] M. Cherrington, F. Thabtah, J. Lu, and Q. Xu, "Feature Selection: Filter Methods Performance Challenges", 2019 International Conference on Computer and Information Sciences (ICCIS), 2019. doi:10.1109/iccisci.2019.8716478

[42] S. Valle, W. Li, and S.J. Qin, "Selection of the Number of Principal Components: The Variance of the Reconstruction Error Criterion with a Comparison to Other Methods", Industrial & Engineering Chemistry Research, 38(11), pp 4389–4401, 1999. doi:10.1021/ie990110i

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 17548 - 17558
Submitted 20 October 2020; Revised 20 November 2020; Accepted 14 December 2020.

Consensus Sequence Characterization Of Rnp-1 And Rnp-2 Along With Base Stacking Interactions With Ribonucleuotide

Isha Sehrawat[1] and Yasha Hasija[2]
[1]Department of Biotechnology, Delhi Technological University, Delhi, INDIA
[2]International Centre for Genetic Engineering and Biotechnology, Delhi, INDIA

ABSTRACT

RNA recognition motif or RRM is the most abundantly present nucleic acid binding domain in eukaryotes. It is also known as Ribonucleoprotein/ RNA binding domain, and typically consists of about 90 amino acids. RRM fold typically comprises of 2 alpha helices and 4 antiparallel beta strands arranged as β1α1β2β3α2β4. This recognition motif consists of two consensus sequences, RNP1 and RNP2 with eight and sixth amino acids, respectively. RNP1 and RNP-2 are conserved sequences with their amino acid constitution being RNP1 with sequence in single letter code predicted to be as K/R-G-F/Y-G/A-F/Y-V/I/L-X-F/Y and RNP2 whose sequence in single letter code is V/I/L-F/Y-V/I/L-X-N/L. RNP-1 is centrally located while RNP-2 is located at the N terminal. It is observed that RRM domain of these proteins interact with nucleic acids through base stacking of aromatic rings. The base stacking interactions observed is mainly by the second amino acid of RNP-2 and by third, fifth and/or eighth amino acid of RNP-1. In this paper we show the frequency of the occurrence of aromatic amino acids at the specific locations like 2nd position in RNP-2 and 3rd, 5th and 8th position in RNP-1. We have also characterized the stacking interactions of these bases with nucleic acids by measuring the distance between the amino acid residues of the proteins and the nucleic acid bases, and also characterized the type of stacking interactions observed (Parallel or T-shaped pi-pi interactions)

Keywords:
Consensus sequences, Protein, RNP, RRM, stacking interactions

1. Introduction

Proteins have specialized structural folds that perform very specific functions like enzymatic reactions, catalysis etc. These folds or structures are called as domains. Domains are self-stabilizing structure that folds independently from the rest of protein structure [20]. There can be more than one domain in a protein with different functions similarly one domain may be present in more than one protein. Domains are not protein specific; they are characteristic folds of amino acid sequence capable of performing a specialized function. Examples of domains:- zinc finger domain, EF hand domain, KH-domain, RRM. It is also hypothesized that multi-domain proteins would have emerged from single domain protein by domain insertion[20].

Protein RNA interactions are very important to a cell, metabolically[25]. Its various roles constitute:- polyadenylation, splicing, stability of messenger RNA etc.[25] one of such interactions is mediated by a domain called RNA recognition motif (short form:- RRM) is a nucleic acid binding domain present in ample amounts in eukaryotes and is present throughout all the diverse kingdoms of life, although in fewer abundance than eukaryotes[21]. Irrespective of its name, this domain has been shown to exhibit interactions with diverse classes of substrates like DNA, lipids and other proteins[16, 15]. RRM occurs in over 1% of all human proteins[6]. This motif is characterized by a specific arrangement of alpha helices and beta strands as $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$[11]. This domain interacts with RNA and other substrates by a characteristic consensus sequence, which are:- RNP1 written in 3 letter code for amino acid as (Lys/Arg-Gly-Phe/Tyr-Gly/Ala-Phe/Tyr-Val/Ile/Leu-X-Phe/Tyr) and RNP2 also in three letter code for amino acids (Ile/Val/Leu-Phe/Tyr-Ile/Val/Leu-X-Asn-Leu)[15]. These consensus sequences are found on beta strand 3 and 1, respectively[15]. This motif binds a diverse group of RNA with each protein being very specific to

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 17548 - 17558
Submitted 20 October 2020; Revised 20 November 2020; Accepted 14 December 2020.

the type of RNA bound. A domain with such wide prevalence across different classification points to a very important role in cell dynamics[21].

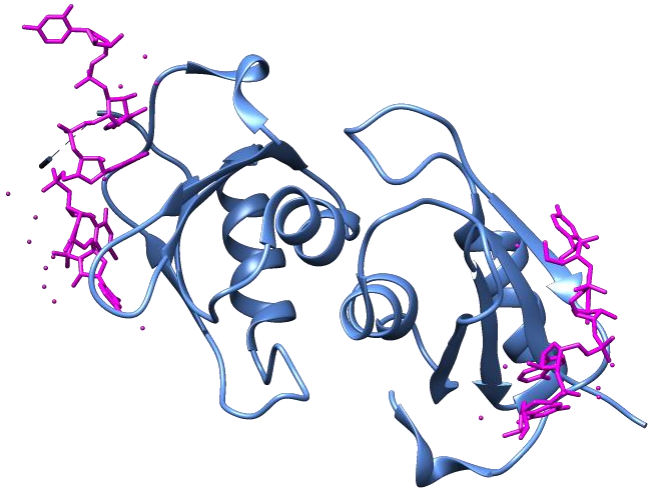


Figure 1- **X**-ray structure of human RBPMS in complex with the RNA[24] (PDB ID- 5DET) where metallic blue colour ribbon is protein chain and magenta pink color in form of line structure is for RNA chain.

How was this domain recognized? This domain was first identified in 1980s when it was seen that almost always pre-mRNAs and heterogeneous nuclear RNA forms a complex with proteins[21]. The consensus sequence of 90 amino acids was found due to detailed examination of PABP and HnRNP protein C[21].

The RRM domain is capable of binding at most to 7-8 nucleotides; but it usually interacts with only 2-4 nucleotides with lower affinity[15]. This domain has shown various ways in which it achieves this specificity towards binding. The easiest way to increase specificity is to bind to the longest RNA possible[15]. This method is employed by many proteins; one of them is FOX-1 protein which increases its surface area for attachment of proteins exponentially by making helix 1 and 2 indulge in binding interactions of protein. Second method employed by proteins is the presence of multiple RRM domains for example in PTB, nPTB, PRP24[15]. PRP24 deserves an important mention in this scenario; this protein has 3 RRM domains which encircle a RNA molecule forming a ring like structure[15].

After characterizing of this domain, there has been an exponential increase in the number of resolved structures for RRM in the past 20 years[15]. This high input of protein structures has led to a lot of findings to come forth regarding this structure. Although the domain has been hypothesized to have a definite topology, this can vary among the various classes of proteins. Some proteins have shown to have an extra beta strand or extra alpha helix. And in many RNA binding proteins this domain is not present, is not involved in any contacts or the binding surfaces are not used[15]. These were either due to masking of beta sheet interaction by an alpha helical structure like in Tetrahymena p65 protein[17]or because of pseudoRRM domains interacting with the substrates[15]. A combination of multiple RRM domains helps solving 2 purposes: - one, it helps in recognition of long nucleotides; two, it increases the binding affinity drastically[21]. Although, some proteins also provide evidence against this by having multiple RRMs which do not interact with each other and also bind to the RNA substrate independently of each other[22].

The N and C- terminal of a protein is very important in RNA-protein interactions to enhance their binding affinity by stacking internally to give a more stable and energetically favorable

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 17548 - 17558
Submitted 20 October 2020; Revised 20 November 2020; Accepted 14 December 2020.

condition[21]. The c-terminal of polypeptide chain helps in the differentiation of DNA from RNA by interacting with the 2'OH group of sugar[21]. The c-terminus is also seen to inhibit binding by making salt bridges with the key positions of RNP and thus preventing any interactions.

RRM and RNA interact through various methods; one of them is base stacking interactions. ∏-∏ interactions are established between a nitrogenous ring and an aromatic ring which contains a pi- electron. Amino acids containing aromatic rings are Trp, His, Phe, and Tyr and the charged amino acids Arg, Glu, and Asp also have an aromatic ring[6]. Pi-pi parallel stacking interactions occur most frequently with the distance of 2.7 to 4.3 Å between the atoms.

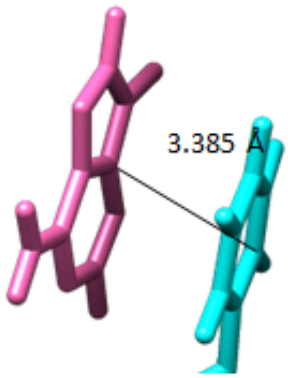


Figure 2. RBMY protein (PDB ID-2FY1)[23] showing parallel pi-pi stacking of F-52 in light blue with A-121 in pink.

Another stacking interaction is t-shaped stacking of the bases; here one of the rings is stacked perpendicularly to the other[6]. These stacking interactions are present widely in the RRM-RNA complex. Other than this, there are hydrophobic interactions, hydrogen- bonding, and more which are out of the scope of this paper[6].

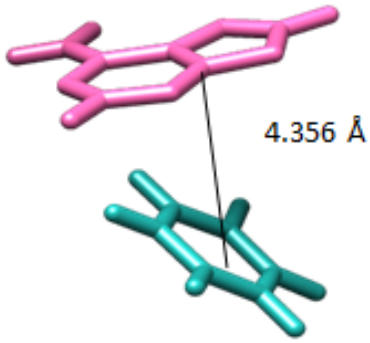


Figure 3: RBMY protein (PDB ID-2FY1)[23] showing T-shaped stacking of F-11 in green with A-120 in pink.

About 200 protein structures were visualized using chimera. Distances between the aromatic rings were calculated by first defining a centroid for respective rings using structure measurement function of chimera. The distance was then calculated between these centroids using command line action and the command "distance c1, c2" where c1 and c2 are the supposed names for the centroids. This will give us a calculated distance in angstrom at the bottom of the chimera window.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 17548 - 17558
Submitted 20 October 2020; Revised 20 November 2020; Accepted 14 December 2020.

Figures and tables:

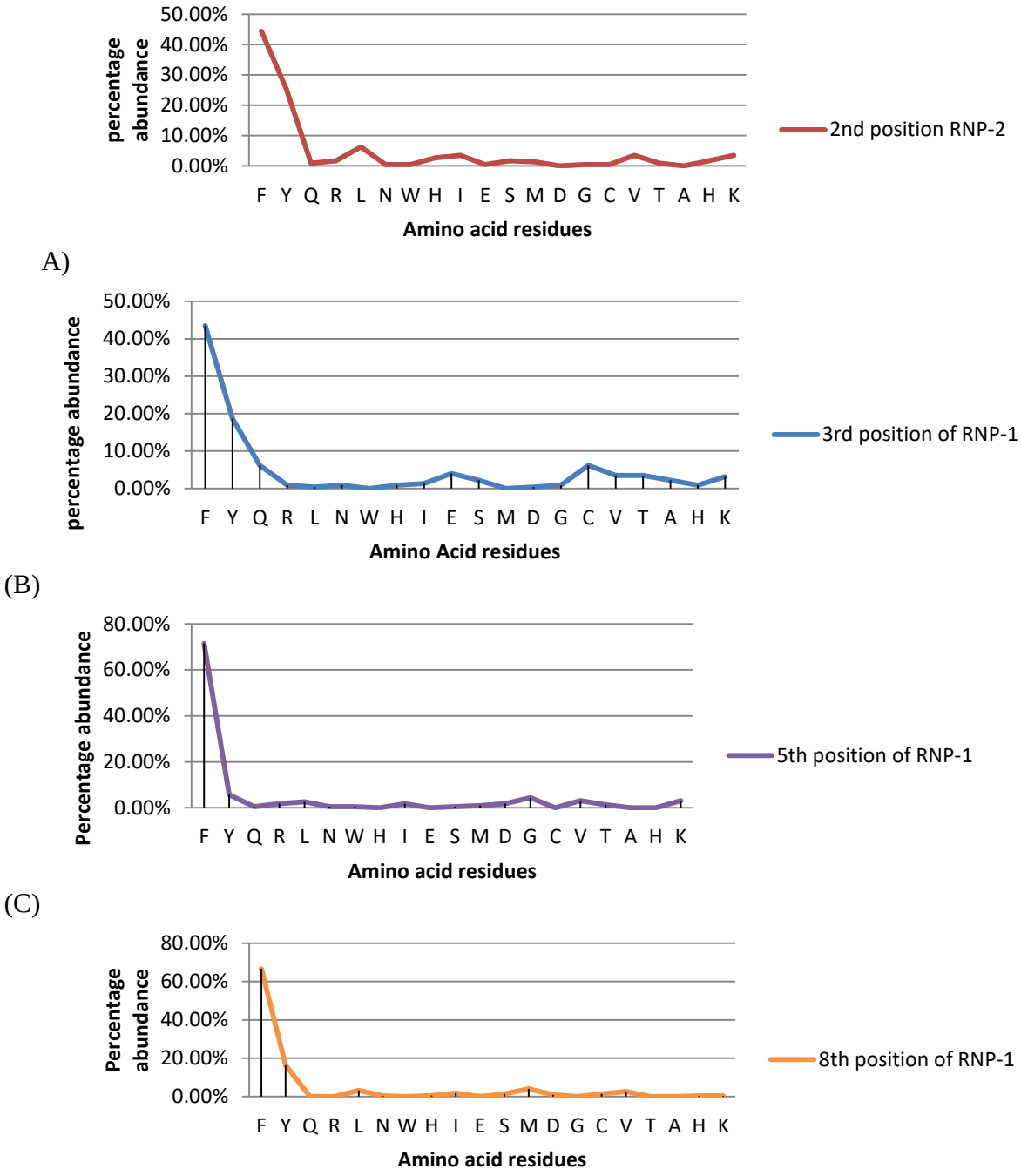
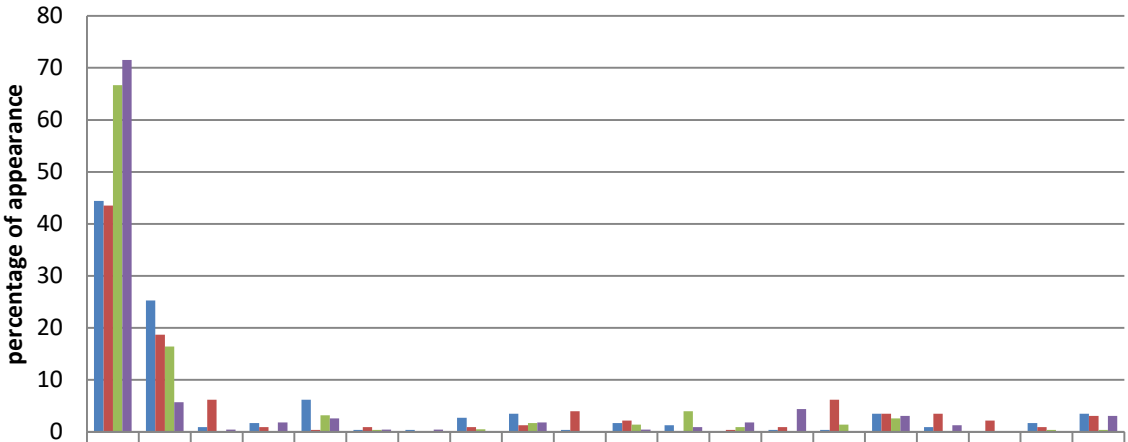


FIGURE 4: Relative Abundance In Percentage Of 20 Amino Acid Residues At The (A) Second Position Of RNP-2, And (B) Third Position, (C) Fifth Position And (D) Eighth Position Of RNP-1 In 198 Structures.

Table 1 The Bar Graph Shows The Abundance Of Different Amino Acid Residues In Protein Structures Complexed With And Without RNA At The Respective Positions (Blue For 2nd Position Of RNP-2, Red For 3rd Position Of RNP-1, Green For 8th Position Of RNP-1 And

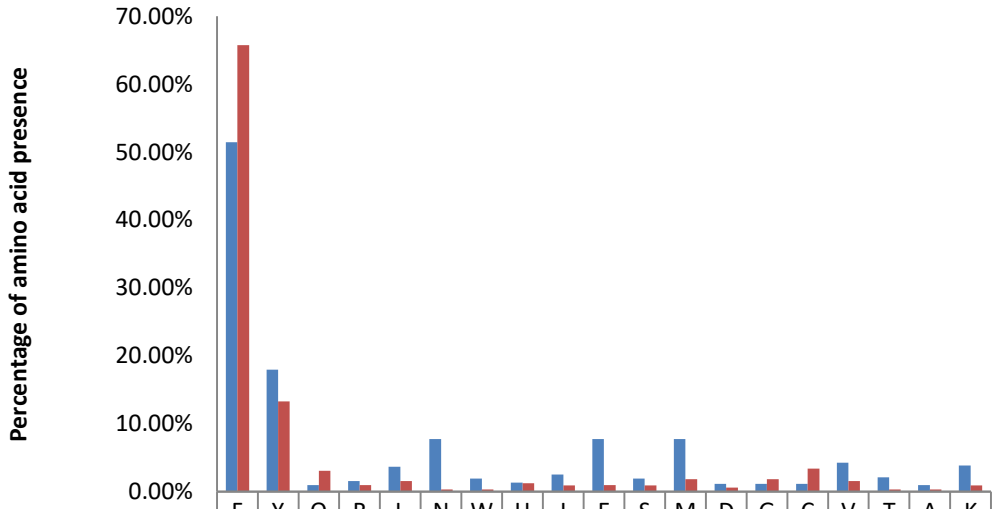
Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 17548 - 17558
Submitted 20 October 2020; Revised 20 November 2020; Accepted 14 December 2020.

Purple For 5th Position Of RNP-1). The Abundances Shown Are In Percentage, Calculated By Taking The Sum Of Each Individual Amino Acid And Then Diving It By The Total Number Of Residues At Each Position Which Was Then Multiplied By Hundred. The Amino Acids Are Written Down In Single Letter Code**.**



	F	Y	Q	R	L	N	W	H	I	E	S	M	D	G	C	V	T	A	H	K
2nd	44	25	0.9	1.7	6.2	0.4	0.4	2.7	3.5	0.4	1.7	1.3	0	0.4	0.4	3.5	0.9	0	1.7	3.5
3rd	44	19	6.2	0.9	0.4	0.9	0	0.9	1.3	4	2.2	0	0.4	0.9	6.2	3.5	3.5	2.2	0.9	3.1
8th	67	16	0	0	3.2	0.4	0	0.5	1.7	0	1.4	4	0.9	0	1.4	2.6	0	0	0.4	0.4
5th	72	5.7	0.4	1.8	2.6	0.4	0.4	0	1.8	0	0.4	0.9	1.8	4.4	0	3.1	1.3	0	0	3.1

Table 2 The graph shows the percentage of each residue present collectively at the 4 sites (2nd position of RNP-2, 3rd, 5th and 8th position of RNP-1) in the structures studied. The percentage was calculated by taking the sum of each amino acid residue shown in single letter code, present at the 4 sites, dividing it by the total number of residues of all sites and multiplying it by hundred.



	F	Y	Q	R	L	N	W	H	I	E	S	M	D	G	C	V	T	A	K
Without RNA	51.	17.	0.9	1.5	3.6	7.7	1.9	1.3	2.5	7.7	1.9	7.7	1.1	1.1	1.1	4.2	2.1	0.9	3.8
RNA complexed	65.	13.	3.0	0.9	1.5	0.3	0.3	1.2	0.9	0.9	0.9	1.8	0.6	1.8	3.3	1.5	0.3	0.3	0.9

Table 3 the pie chart shows the calculated percentage of the amino acid residues showing different types of stacking interactions (parallel pi-pi or T-shaped) with different bases. Total

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 17548 - 17558
Submitted 20 October 2020; Revised 20 November 2020; Accepted 14 December 2020.

number of each interaction was calculated and then divided by the sum of total interactions observed. The result was then multiplied by 100.

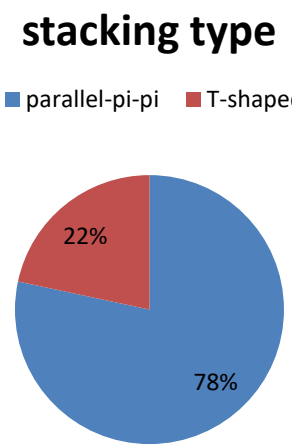
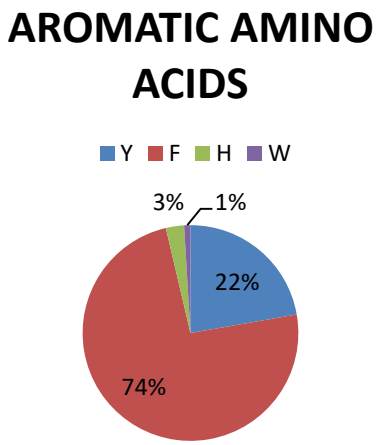
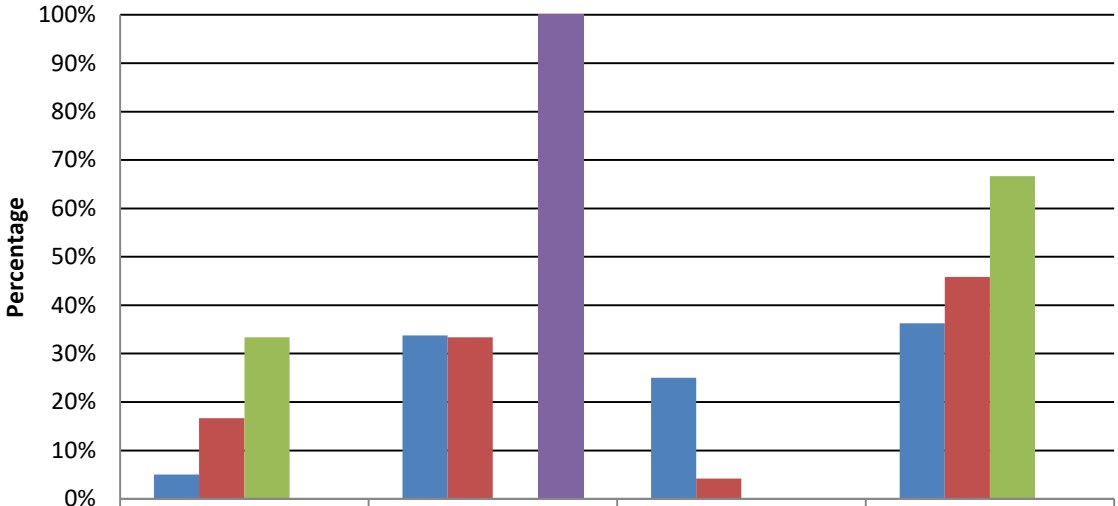


Table 4 the pie chart shows the percentage in which different aromatic amino acids are present for stacking with different bases



Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 17548 - 17558
Submitted 20 October 2020; Revised 20 November 2020; Accepted 14 December 2020.



	C	A	G	U
F	5%	33.75%	25%	36.25%
Y	16.67%	33.34%	4.16%	45.83%
H	33.34%	0%	0%	66.67%
W	0%	100%	0%	0%

TABLE 5 the above graph shows the amount of stacking interaction observed by the aromatic amino acids (phenylalanine as F, tyrosine as Y, histidine as H and tryptophan as W) with the four bases present in a ribonucleic acid (cysteine as C, adenosine as A, guanine as G and uracil as U) in percentages irrespective of their position on the two different RNP.

2.Results:

COMPLIANCE OF DIFFERENT PROTEINS TO THE PREDICTED SEQUENCE

It was also seen that between the two RNP sequences it is seen that RNP-1 was adhering to the proposed sequence and the faltering sequences also didn't differ too much, while there were many variations seen in RNP-2 and some changes were very drastic. The changes seen irrespective of the key positions talked about in the paper for stacking interactions were also very drastic.

DIFFERENT PERCENTAGES OBSERVED IN RNP-1 AND RNP-2 SEQUENCE

After comparing about 200 structures, it can be concluded that Phenyl-alanine(F) is the most abundant in the 2nd position of RNP-2 and in 3rd, 5th and 8th position of RNP-1in the percentages of 44%, 44%, 72% and 67%, respectively. Tyrosine(Y) took the second lead after Phenylalanine, 25%, 19%, 5.7% and 16% at 2nd position of RNP-2 and 3rd, 5th and 8th position, respectively (FIGURE 4, TABLE 1). Although there is a vast difference between the percentages of F and Y but the dominance of rest of the amino acids exhibited still higher differences. The collective percentages of all the amino acids in the different positions added up shows even clear demarcation, 51% of F in the protein structures that didn't have RNA complexed with them and 65% in the proteins complexed with RNA; 17% of Y in the protein structures that didn't have RNA complexed with them and 13% in the proteins complexed with RNA (TABLE 2).

ABUNDANCE OF AROMATIC AMINO ACIDS AMONG THE STRUCTURES VISUALISED

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 17548 - 17558
Submitted 20 October 2020; Revised 20 November 2020; Accepted 14 December 2020.

In the structures with RNA bound proteins, it was observed that Phenylalanine and tyrosine were the amino acids with the highest percentage 74% and 22%, respectively (TABLE 4). The stacking interactions were majorly parallel pi-pi interactions with a very distinct majority, 78%, over T-shaped stacking, 22% (TABLE 3). Phenylalanine was bound mostly to uridine, although the percentage of uridine bound and adenosine bound protein was very close, uridine was about 3% more than adenosine. Tyrosine also seemed to favor uridine interactions, rather more fiercely than Phenylalanine since the difference between the percentages of other bases interaction was very high (TABLE 5).

3Conclusion And Discussion:

The RRM domain however constricted due to the presence of consensus sequence, have a variety of functions. This can be seen in every species; like in Homo sapiens we saw variety of different proteins, having the consensus sequence. The proposed RNP sequences for RNP-1 and RNP-2 is conserved throughout the Homo sapien species irrespective of the type of protein. At some key positions, 2^{nd} position of RNP-2, 3^{RD}, 4^{TH} and 5^{TH} position of RNP-1, the highest percentage was of aromatic amino acids as is hypothesized over the years.

In the considered protein sequences, it was observed that the amino acid phenylalanine or phe (in three letter code) or F (in single letter code) was seen to be predominant in all the four places separately and together as well. Its presence was in majority even when the structures complexed with RNA were compared to the ones not complexed with any nucleic acid. It can be said that phe is favored more than the other aromatic amino acids.

After phenylalanine, in every case, the lead was of tyrosine, tyr (as three letter code) or Y as single letter code. They both were also seen to be consistent in favoring one nucleic acid i.e., uridine or U as single letter code. The second came adenine, A as single letter code, whereas the difference between their abundance in stacking presence was merely 2.5% but comparatively the difference between U and A with other nucleic acids were very high. Cytosine was the lowest to be stacked with by any of the aromatic bases found.

In conclusion, it is safe to say that the proposed RNP sequences for RNP-1 and RNP-2 is conserved throughout the *Homo sapiens* species irrespective of the type of protein as the highest percentage was of aromatic amino acid in the positions: 2^{nd} position of RNP-2, 3^{rd}, 5^{th} and 8^{th} position of RNP-1. The highest percentage over-all was of phenylalanine and the amino acid favored uridine more than rest of the bases. The favoring was consistent with tyrosine as well. Cytosine was the lowest to be stacked with by any of the aromatic bases found.

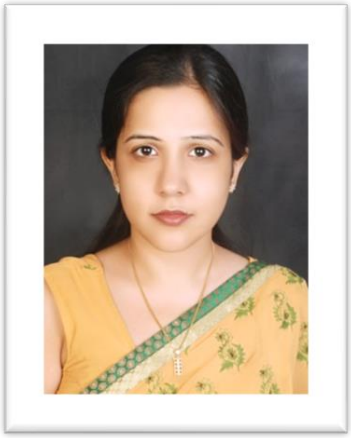
Acknowledgement
Authors thank Dr. Neel Sarovar Bhavesh and Dr. Priyanka, ICGEB for their scientific inputs.

Authors Profile

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 17548 - 17558
Submitted 20 October 2020; Revised 20 November 2020; Accepted 14 December 2020.



Isha Sehrawat is a student at Delhi technological university pursuing her Masters in biotechnology. She did her bachelors from amity university Haryana. she did her summer internship from ICGEB, new delhi, under Dr. Neel Sarovar Bhavesh who is the PI of transcription regulation team.



Dr Yasha Hasija is currently working as Associate Professor, Department of Biotechnology and Associate Dean (Alumni Affairs) at Delhi Technological University. She has published more than 50 research articles and review papers in national and international journals of high repute, 02 books; 10 book chapters; and 26 conference papers. She is also on the Editorial Board of numerous international journals. Dr. Hasija's work has earned recognition and received several prestigious awards, including the Govt. of India- Department of Science and Technology Award for attending the meeting of Nobel Laureates and Students in Lindau, Germany in 2002; and Human Gene Nomenclature Award at the Human Genome Meeting-2010 held at Montpellier, France. She has also been awarded Commendable Research Award, DTU for three consecutive years- 2018, 2019 and 2020, and Premier Research Award, DTU in 2021.
Dr Hasija is the Project Investigator of several sponsored research projects from Govt. of India departments including DST, CSIR and DBT. She has delivered more than 20 invited talks at several prestigious universities and institutions. She is an active researcher who has supervised 4 Ph.D. awardees and 40 M.Tech students at DTU.
Her broad areas of research include genome informatics, integration of genome-scale data for systems biology, and machine learning applications in healthcare

References

[1] Landsman, David. (1992). RNP-1, an RNA-binding motif is conserved in the DNA-binding cold shock domain. Nucleic acids research. 20. 2861-4. 10.1093/nar/20.11.2861.

[2] Daubner GM, Cléry A, Allain FH. RRM-RNA recognition: NMR or crystallography…and new findings. Curr Opin Struct Biol. 2013 Feb;23(1):100-8. doi: 10.1016/j.sbi.2012.11.006. Epub 2012 Dec 14. PMID: 23253355.

[3] Dasti A, Cid-Samper F, Bechara E, Tartaglia GG. RNA-centric approaches to study RNA-protein interactions in vitro and in silico. Methods. 2020 Jun 1;178:11-18. doi: 10.1016/j.ymeth.2019.09.011. Epub 2019 Sep 26. PMID: 31563541.

[4] Aaron R. Poole, Ian Vicino, Hiro Adachi, Yi-Tao Yu, Michael D. Hebert Regulatory RNPs: A novel class of ribonucleoproteins that contribute to ribosome heterogeneity doi: https://doi.org/10.1101/128090

[5] Jayant Udgaonkar and Susan Marqusee Folding and binding https://doi.org/10.1016/j.sbi.2013.01.002

[6] Meredith Corley,1 Margaret C. Burns,1,2 and Gene W. Yeo1,2,3, *How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms https://doi.org/10.1016/j.molcel.2020.03.011

[7] Ramanathan M, Porter DF, Khavari PA. Methods to study RNA-protein interactions. Nat Methods. 2019 Mar;16(3):225-234. doi: 10.1038/s41592-019-0330-1. Epub 2019 Feb 25. Erratum in: Nat Methods. 2019 Apr;16(4):351. PMID: 30804549; PMCID: PMC6692137.

[8] Bleichert F, Baserga SJ. Ribonucleoprotein multimers and their functions. Crit Rev Biochem Mol Biol. 2010 Oct;45(5):331-50. doi: 10.3109/10409238.2010.496772. PMID: 20572804; PMCID: PMC2939948.

[9] Lee SR, Lykke-Andersen J. Emerging roles for ribonucleoprotein modification and remodeling in controlling RNA fate. Trends Cell Biol. 2013 Oct;23(10):504-10. doi: 10.1016/j.tcb.2013.05.001. Epub 2013 Jun 4. PMID: 23756094; PMCID: PMC3983964.

[10] Thirumalai D, Liu Z, O'Brien EP, Reddy G. Protein folding: from theory to practice. Curr Opin Struct Biol. 2013 Feb;23(1):22-9. doi: 10.1016/j.sbi.2012.11.010. Epub 2012 Dec 21. PMID: 23266001.

[11] Cléry A, Blatter M, Allain FH. RNA recognition motifs: boring? Not quite. Curr Opin Struct Biol. 2008 Jun;18(3):290-8. doi: 10.1016/j.sbi.2008.04.002. PMID: 18515081.

[12] Al-Hashimi HM, Walter NG. RNA dynamics: it is about time. Curr Opin Struct Biol. 2008 Jun;18(3):321-9. doi: 10.1016/j.sbi.2008.04.004. Epub 2008 Jun 9. PMID: 18547802; PMCID: PMC2580758.

[13] Hafner, M., Katsantoni, M., Köster, T. et al. CLIP and complementary methods. Nat Rev Methods Primers 1, 20 (2021). https://doi.org/10.1038/s43586-021-00018-1

[14] Auweter SD, Oberstrass FC, Allain FH. Sequence-specific binding of single-stranded RNA: is there a code for recognition? Nucleic Acids Res. 2006;34(17):4943-59. doi: 10.1093/nar/gkl620. Epub 2006 Sep 18. PMID: 16982642; PMCID: PMC1635273.

[15] Afroz T, Cienikova Z, Cléry A, Allain FHT. One, Two, Three, Four! How Multiple RRMs Read the Genome Sequence. Methods Enzymol. 2015;558:235-278. doi: 10.1016/bs.mie.2015.01.015. Epub 2015 Mar 12. PMID: 26068744.

[16] Clingman, C. C., Deveau, L. M., Hay, S. A., Genga, R. M., Shandilya, S. M. D., Massi, F., et al. (2014). Allosteric inhibition of a stem cell RNA-binding protein by an intermediary metabolite. eLife, 3, 1–26. http://dx.doi.org/10.7554/Elife.02848.

Cloonan, N., Forrest, A. R. R., Kolle, G., Gardiner, B. B. A., Faulkner, G. J., Brown, M. K., et al. (2008). Stem cell transcriptome profiling via massive-scale mRNA sequencing. Nature Methods, 5(7), 613–619.

[17] Singh, M., Choi, C. P., & Feigon, J. (2013). XRRM: A new class of RRM found in the telomerase La family protein p65. RNA Biology, 10(3), 353–359. http://dx.doi.org/10.4161/rna.23608.

[18] Kent SB (February 2009). "Total chemical synthesis of proteins". Chemical Society Reviews. 38 (2): 338–51. doi:10.1039/b700141j. PMID 19169452

[19] Van Holde KE, Mathews CK (1996). Biochemistry. Menlo Park, California: Benjamin/Cummings Pub. Co., Inc. ISBN 978-0-8053-3931-4.

[20] Xu D, Nussinov R (1 February 1998). "Favorable domain size in proteins". Folding & Design. 3 (1): 11–7. doi:10.1016/S1359-0278(98)00004-2. PMID 9502316.

[21] The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression Christophe Maris Cyril Dominguez Frédéric H.-T. Allain https://doi.org/10.1111/j.1742-4658.2005.04653.x

[22] Cléry A, H.-T. Allain F. FROM STRUCTURE TO FUNCTION OF RNA BINDING DOMAINS. In: Madame Curie Bioscience Database [Internet]. Austin (TX): Landes Bioscience; 2000-2013. Available from: https://www.ncbi.nlm.nih.gov/books/NBK63528/

[23] Skrisovska L, Bourgeois CF, Stefl R, Grellscheid SN, Kister L, Wenter P, Elliott DJ, Stevenin J, Allain FH. The testis-specific human protein RBMY recognizes RNA through a novel mode of interaction. EMBO Rep. 2007 Apr;8(4):372-9. doi: 10.1038/sj.embor.7400910. Epub 2007 Feb 23. PMID: 17318228; PMCID: PMC1852761.

[24] Teplova M, Farazi TA, Tuschl T, Patel DJ. Structural basis underlying CAC RNA recognition by the RRM domain of dimeric RNA-binding protein RBPMS. Q Rev Biophys. 2016 Jan;49:e1. doi: 10.1017/S0033583515000207. Epub 2015 Sep 8. PMID: 26347403; PMCID: PMC4783296.

[25] Chen, Y. et al. “Identifying RNA-binding residues based on evolutionary conserved structural and energetic features.” Nucleic Acids Research 42 (2014): e15 - e15.

[26] Rao JN, Schweimer K, Wenzel S, Wöhrl BM, Rösch P. NELF-E RRM undergoes major structural changes in flexible protein regions on target RNA binding. Biochemistry. 2008 Mar 25;47(12):3756-61. doi: 10.1021/bi702429m. Epub 2008 Feb 28. PMID: 18303858.

2021 2nd International Conference for Emerging Technology (INCET)
Belgaum, India. May 21-23, 2021

COVID 19, Pneumonia and Other Disease Classification using Chest X-Ray images

Abhishek Panwar
Department of Computer Science & Engineering
Delhi Technological University Shahbad
Daulatpur, Main Bawana Road, DL - 42, IN
abhishekpanwar971@gmail.com

Akshay Dagar
Department of Computer Science & Engineering
Delhi Technological University Shahbad
Daulatpur, Main Bawana Road, DL - 42, IN
akshaydagar98@gmail.com

Vishal Pal
Department of Computer Science & Engineering
Delhi Technological University Shahbad
Daulatpur, Main Bawana Road, DL - 42, IN
vishalpal1999@gmail.com

Dr. Vinod Kumar
Department of Computer Science & Engineering
Delhi Technological University Shahbad
Daulatpur, Main Bawana Road, DL - 42, IN
vinod_k@dtu.ac.in

2021 2nd International Conference for Emerging Technology (INCET) | 978-1-7281-7029-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/INCET51464.2021.9456192

***Abstract*— In this report, we propose an alternative way of detecting the Covid-19 disease using Convolutional Neural Networks based deep learning models which will classify between Covid-19 and other similar respiratory diseases such as other types of Viral Pneumonia or Bacterial Pneumonia, using images of Chest X-Rays. The models developed as part of this project achieved approximately 98% accuracy on a test dataset of 300 images.**

Keywords—Covid-19, Pneumonia, Disease Classification

I. INTRODUCTION

Coronavirus disease 2019 (COVID-19) is an infectious disease caused by severe acute respiratory syndrome Corona virus 2 (SARS-CoV-2). It was first identified in December 2019 in Wuhan, Hubei, China, and has resulted in an ongoing pandemic. As per WHO on 21 March 2021, more than 123 million cases have been reported across 223 countries and territories with more than 2.7 million deaths: more than 69.5 million people have recovered Common symptoms include fever, cough, fatigue, shortness of breath or breathing difficulties, and loss of smell and taste. While most people have mild symptoms, some people develop acute respiratory distress syndrome (ARDS) possibly precipitated by cytokine storm, multi-organ failure, septic shock, and blood clots. The incubation period may range from one to fourteen days. This disease generally spreads through air during close proximity of an infected person, contaminated surfaces etc. Currently the standard method of diagnosis is by real-time reverse transcription polymerase chain reaction (rRT-PCR) from a nasopharyngeal swab. The Chest X-ray imaging may also be helpful for diagnosis in individuals where there is a high suspicion of infection based on symptoms and risk factors, however guidelines do not recommend using it for routine screening currently. In our study here, we trained our models on openly available chest X-ray dataset maintained by physicians and sources like Kaggle. To avoid any sort of overfitting due to shortage of data, we used various renowned image augmentation techniques. Our models had shown high accuracy on Testing dataset. However , we need more data and extensive testing to put it in production. In coming section, we will discuss about our literature overview, our methodology and results.

II. LITERATURE OVERVIEW

Mucahid Barstugan et. al. [5] study presents early phase detection of Coronavirus (COVID-19), which is named by the World Health Organization (WHO), by machine learning methods. The detection process was implemented on abdominal Computed Tomography (CT) images. The expert radiologists detected from CT images that COVID-19 shows different behaviours from other viral pneumonia. Therefore, the clinical experts specify that COVİD-19 virus needs to be diagnosed in early phase. For detection of the COVID-19, four different datasets were formed by taking patches sized as 16x16, 32x32, 48x48, 64x64 from 150 CT images. The feature extraction process was applied to patches to increase the classification performance. Grey Level Co-occurrence Matrix (GLCM), Local Directional Pattern (LDP), Grey Level Run Length Matrix (GLRLM), Grey-Level Size Zone Matrix (GLSZM), and Discrete Wavelet Transform (DWT) algorithms were used as feature extraction methods. Support Vector Machines (SVM) classified the extracted features. 2-fold, 5-fold and 10-fold cross-validations were implemented during the classification process. Sensitivity, specificity, accuracy, precision, and F-score metrics were used to evaluate the classification performance.

In Enzo Tartaglione et. al. [8] insights are provided and also warnings are raised on what is reasonable to expect by applying deep learning to COVID classification of CXR images. A methodological guide and critical reading of an extensive set of statistical results that can be obtained using currently available datasets is provided. In particular,the challenge posed by current small size COVID data and show how significant can be the bias introduced by transfer-learning using larger public non-COVID CXR datasets is taken. Contribution is also made by providing results on a medium size COVID CXR dataset, just collected by one of the major emergency hospitals in Northern Italy during the

© IEEE 2021. This article is free to access and download, along with rights for full text and data mining, re-use and analysis.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:56:59 UTC from IEEE Xplore. Restrictions apply.

peak of the COVID pandemic. These novel data allow for the contribution to validate the generalization capacity of preliminary results circulating in the scientific community. The conclusions shed some light into the possibility to effectively discriminate COVID using CXR.

However, in Gianluca Maguolo et. al. paper [9], different testing protocols used for automatic COVID-19 diagnosis from X-Ray images in the recent literature are compared and evaluated. It is shown that similar results can be obtained using X-Ray images that do not contain most of the lungs. It is possible to remove the lungs from the images by turning to black the center of the X-Ray scan and training their classifiers only on the outer part of the images. Hence, it is deduced that several testing protocols for the recognition are not fair and that the neural networks are learning patterns in the dataset that are not correlated to the presence of COVID-19. Finally, it is shown that creating a fair testing protocol is a challenging task, and they provide a method to measure how fair a specific testing protocol is. In the future research it is suggested to check the fairness of a testing protocol using their tools and they encourage researchers to look for better techniques than the ones that they propose.

This paper tries to incorporate better strategies in order to overcome aforementioned shortcomings. However, lack of ample sized dataset acts as a bottleneck in obtaining more accurate results.

III. METHODOLOGY

In this section, we have divided our work in four parts. Firstly, we will discuss about the collection of data, data sources and its relevance. Then, we will move to data processing and preparation part where we discuss about how we prepared the data for training. Then, we will see the different model architectures that we used for training the parameters and finally we will see the training and testing of the data for various parameters.

The workflow of our work throughout this process, can be summarised as shown in the figure down below.

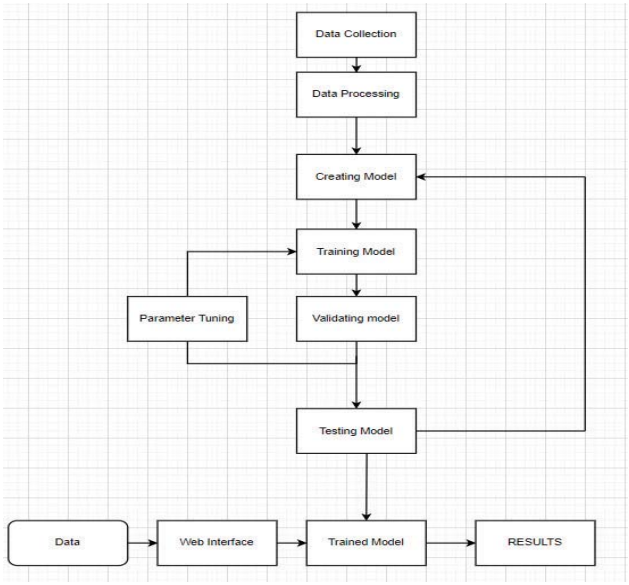


Fig. 1. Workflow Diagram

A. Data Collection

Data is a very important input for any machine learning or deep learning problem. A relevant collection of data helps in analysing a machine learning problem better. In our problem set, we have collected data majorly from two different and prominent sources present online. Firstly, we have data from Covid 19 chest X-ray dataset repository on Github maintained by Joseph Paul Cohen, Paul Morrison, Lan Dao, Karsten Roth, Tim Q Duong and Marzyeh Ghassemi. The data of this repository is collected from various hospitals and physicians. This data was collected from public sources like hospitals and maintained under the project approved by University of Montreal's Ethics Committee. We have chosen the 'AP' and 'PA' view images of chest X-ray for our identification. We further classified it in majorly three classes as covid affected, non-covid pneumonia affected and normal/ other disease affected. Further, we collected the data of Pneumonia and other diseases from Kaggle competition. Apart from these sources, we supplemented our data set from various online sources and studies done before. Finally, after combining these datasets, we were able to gather a total of 4563 images (478 Covid 19 or Covid 19 affected pneumonia images, 2647 other pneumonia images and the rest 1438 were the normal/ other disease images).

B. Data Preparation and Processing

From this combined dataset, we will create a testing dataset containing 100 images belonging to each of the 3 classes (Covid-19 Pneumonia images, Other Pneumonia Images, and Other Diseases/No disease images) and validation dataset containing 78 images belonging to each of the three classes. This will leave us with 300 Covid-19 images, 2469 other Pneumonia images and 1260 images belonging to other diseases/no diseases class

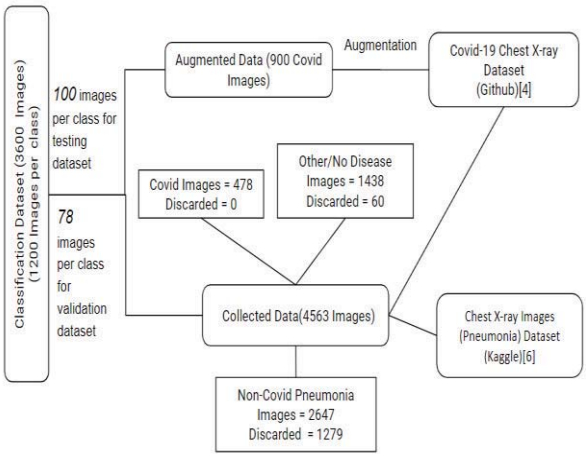


Fig. 2. Dataset Preparation

The leftover training dataset is highly imbalanced and if we use this dataset to train our models, there are high chances of overfitting as the model will be able to obtain a high accuracy by predicting the majority class(es). To balance the collected dataset, we used Image Augmentation techniques. The different augmentation techniques we used are as follows:

- Random Rotations
- Random Shifting (left, right, up, or down)

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:56:59 UTC from IEEE Xplore. Restrictions apply.

- Random Gaussian Noise Adding
- Horizontal Flipping (mirror image)
- Random Cropping
- Random Shearing
- Random Scaling

For each Covid-19 Pneumonia Image, we will randomly select 3 of these 10 augmentation functions and apply them to the image to generate 3 augmented copies of the image and save the copies with the original image in our dataset.

After Augmenting, we have 1200 Covid-19 Pneumonia CXR images. However, the dataset is still imbalanced, and we need to reduce the no. of images of the other 2 classes to 1200 each to form a perfectly balanced dataset. Since the other 2 classes contain multiple images from a few persons, we will keep only 1 image per person thus reducing the no. of images belonging to these 2 classes. After that, we will keep only 1200 images from each class to form our final perfectly balanced training dataset.

IV. Model Architectures and classifiers

In this section, we discuss different model architectures we used. In our work, we mainly use two deep learning models which are 4 Layer CNN (Model 1) and AlexNet [10] (Model 2) Model 1 was further classified on the basis of our dataset. We used Model 1 with the original dataset initially collected once and then we used it with a modified dataset which was supplemented with our augmented dataset.

- Model 1 (4 Layer CNN)
- Model 2 (AlexNet)

The two models were trained with the gathered data as follows:

- Initially, we trained Model 1 with the original dataset. The training was done using Image Data Generators which will read images from the dataset folders and resize them to size (224,224) with 3 channels (RGB) and rescale the images to [0,1]. Data generators were also used for validation images. After training, we only want the model weights which give the best validation accuracy to be saved as our final model weights, hence we use the "Model Checkpoint" callback from the Keras library. We saved the model parameters that result in the best accuracy. Finally, we trained this model for 50 epochs with a default batch size of 32 train images per batch.
- Thereafter, We trained Model 2 with the original dataset. This model is based on AlexNet. We will take the Convolution layers of the AlexNet model and discard the Output layers and Dense (hidden) layers and instead attach our own hidden layers and output layer with Softmax activation to the output of the convolution layers. The complete architecture has already been defined in the previous section on "Model Architectures". The training of this model will be done similar to that of model1. The only exception being that we will fine tune the AlexNet weights while training instead of using randomly initialized weights as done for model 1.
- Finally, we used the same Model 1 however we trained it differently using Image Data generators with augmentation. Basically, we will randomly take images from the training dataset folder and apply random augmentations on them, thus generating augmented versions of those images, on the fly, while training. Note that we will not use the original image, only the augmented ones, while training. Since we have already augmented our original dataset, this will be kind of a double augmentation. Augmenting like this can help us reduce overfitting our model since the dataset contains augmented copies of the original 300 images and hence there will be a lot of similarities. The rest of the training was similar to other training above.

V. Experiment Results and Discussion

In this section, we will see the experiment results collection from testing.

We tested the trained models on the test dataset containing 100 samples for each of the 3 classes, the results are as follows:

TABLE I. No. of images of different classes used for training and testing dataset

Class Name	*No. of Images*	
	Training	*Testing*
Covid 19	1200	100
Non –Covid Pneumonia	1200	100
Normal/Other diseases	1200	100

(Note: 0 = Covid-19 Pneumonia, 1 = Other Diseases/No Disease, 2 = Other Pneumonia)

The results of the three models are as follows:

Finally, we will take the weighted average of the predictions of the 3 models to get final predictions. This weighted prediction function given below is the final classifier we used in our work. The weighted average is taken in the following manner:

Count_Class0 = (No. of Models with majority class prediction = 0)

Count_Class1 = (No. of Models with majority class prediction = 1)

Count_Class2 = (No. of Models with majority class prediction = 2)

Final Prediction = ((P1*Count_ClassP1) + (P2*Count_ClassP2) + (P3*Count_ClassP3)) / (Count_ClassP1 + Count_ClassP2 + Count_ClassP3)

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:56:59 UTC from IEEE Xplore. Restrictions apply.

TABLE II. CLASSIFICATION ACCURACY, PRECISION AND RECALL (IN PERCENTAGES) FOR DIFFERENT DEEP LEARNING MODELS IN COVID-19, PNEUMONIA AND OTHER DISEASES/NORMAL CLASSIFICATION ON TRAINING AND TESTING DATASET

Model Name	***Closs***	**Testing/Training**		
		Accuracy	***Precision***	***Recall***
Model 1	Covid	98.67	98.00	98.00
	Pneumonia	97.00	95.00	95.96
	Normal/other Disease	97.67	97.00	96.04
Model 2	Covid	98.00	98.00	96.07
	Pneumonia	97.67	94.00	98.94
	Normal/Other Disease	98.33	99.00	96.11
Model 3	Covid	98.00	95.00	98.95
	Pneumonia	95.00	88.00	96.70
	Normal/Other Disease	95.67	100.00	88.49
Final Result	Covid	99.00	99.00	98.02
	Pneumonia	98.33	95.00	100.00
	Normal / Other Disease	98.67	100.00	96.15

Thus the overall accuracy for Model 1 is 96.67 %, for Model 2 is 97.00% , for Model 3 is 94.34%.The final results show that our classifier is performing well on a test dataset of 300 images and achieving 98% accuracy with 100% accuracy for “Other Diseases/No disease” class and 99% accuracy for “Covid-19 Pneumonia” class.

Due to lack of data for Covid-19, we cannot test it on a larger dataset, however testing the classifier on a dataset of 1000 images each of classes 1 and 2 and 100 images of class 0 (Covid-19) produced almost similar results with slightly reduced accuracy (97%) which is bound to happen on larger datasets).

VI. CONCLUSION AND FUTURE WORK

AlexNet and 4 layer Convolutional Neural Network have been proved to be really worthy and provided us the commendable results in predicting the Covid 19 cases using only the X ray images. Using them, we are able to achieve almost 98% accuracy which is comparatively higher than many other parallel research work going in the field. But since the data available is quite less right now, we cannot completely rely on these models and their results for actual implementation in the Medical industry. So, to conclude we can say that a more exhaustive availability of the required data is needed.

As future work, we plan to test our classifier on more data as it becomes available so as to produce more reliable results as per the medical industry acceptance standards. Till then, we are also looking forward to training some parallel CNN models on the available data in near future and testing parallel CNN models for the same in order to achieve better results.

REFERENCES

[1] R. M. Pereira, D. Bertolini, D. Bertolini, C. N. S. Jr. and Y. M. Costa, "COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios," *Computer Methods and Programs in Biomedicine,* vol. 194, 2020.

[2] A. I. F. AI, AI2, CZI, MSR, Georgetown, NIH and T. W. House, "COVID-19 Open Research Dataset Challenge (CORD-19)," 2020. [Online]. Available: https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge. [Accessed 2020].

[3] J. P. Cohen, P. Morrison and L. Dao, "COVID-19 Image Data Collection," *arXiv,* p. 4, 25 March 2020. K. Elissa, “Title of paper if known,” unpublished.

[4] J. P. Cohen, "covid-chestxray-dataset," 2020.

[5] S. ozturk, U. Ozkaya and M. Barstugan, "Classification of Coronavirus (COVID-19) from X-rayand CT images using shrunken features," *International Journal of Imaging Systems and Technology,* p. 11, 2020.

[6] Kaggle, "COVID-19 chest xray," 2020.

[7] Wikipedia, "Coronavirus disease 2019," 2019. [Online]. Available: https://en.wikipedia.org/wiki/Coronavirus_disease_2019. [Accessed 2020].

[8] E. Tartaglione, C. A. Barbano, C. Berzovini, M. Calandri, and M. Grangetto, “Unveiling COVID-19 from CHEST X-Ray with Deep Learning: A Hurdles Race with Small Data,” *International Journal of Environmental Research and Public Health*, vol. 17, no. 18, p. 6933, Sep. 2020.

[9] G. Maguolo and L. Nanni, "A Critic Evaluation of Methods for COVID-19 Automatic Detection from X-Ray Images," *ArXiv,* vol. abs/2004.12823, 2020.

[10] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," 25th International Conference on Neural Information Processing Systems, vol. 1, pp. 1097-1105, 2012.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:56:59 UTC from IEEE Xplore. Restrictions apply.

2021 2nd International Conference for Emerging Technology (INCET)
Belgaum, India. May 21-23, 2021

COVID-19 Severity Assessment from Chest X-rays using Attention-based Weakly-Supervised Learning

Kunal Chaturvedi
Department of Applied Mathematics
Delhi Technological University
New Delhi, India
kunalchaturvedi98@gmail.com

Tanmay Kansal
Department of Applied Mathematics
Delhi Technological University
New Delhi, India
tanmaykansal702@gmail.com

Saksham Gupta
Department of Applied Mathematics
Delhi Technological University
New Delhi, India
sakgupta55@gmail.com

Dinesh Kumar Vishwakarma
Department of Information Technology
Delhi Technological University
New Delhi, India
dvishwakarma@gmail.com

Naokant Deo
Department of Applied Mathematics
Delhi Technological University
New Delhi, India
naokantdeo@dce.ac.in

2021 2nd International Conference for Emerging Technology (INCET) | 978-1-7281-7029-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/INCET51464.2021.9456449

***Abstract*— The potential of automated severity assessment of COVID-19 pneumonia is immense due to its ability to facilitate clinical decision-making. It enables efficient escalation or de-escalation of COVID-19 care. In this work, we propose an efficient pipeline based on weakly-supervised learning for severity score prediction. In the first stage, Attention feature fusion (AFF)-ResNet-101 is trained on five large Chest X-ray (CXR) datasets. In the second stage, using transfer learning on 94 posteroanterior (PA) COVID-19 CXR images, refined feature maps are extracted. These feature descriptors are then mapped to severity ratings based on Geographic Extent score and lung opacity output using the linear regression model. Experimental results show the efficacy of our proposed architecture for the severity score prediction.**

Keywords*— *COVID-19, severity assessment, weakly-supervised learning, attention mechanism

I. INTRODUCTION

The novel coronavirus SARS-CoV-2, is a fatal respiratory illness that is highly contagious and can be transmitted through the respiratory droplets or by contact with a contaminated surface [1][2]. Cough, fever, loss of smell, shortness of breath, body aches, lung inflammation [3] are some of the common symptoms of this disease. Inflammation in the lungs, can also cause acute respiratory distress syndrome [4]. Compared to SARS, MERS, and other similar respiratory diseases, the COVID-19 spread rate is enormous [5][6] as it can also be spread by asymptomatic COVID-19 carriers. The pandemic has slowed down the world economy and deeply affected the healthcare infrastructure in all the countries. Early detection of this disease can limit the spread of the virus and help in saving lives by facilitating timely treatment.

COVID-19 disease diagnosis requires detection of SARS-CoV-2 through laboratory tests using CT scans, Chest X-rays (CXRs), RT-PCR testing, etc. RT-PCR test is less favourable due to its limited sensitivity, longer detection time, and inconsistent results. Also, due to the limited availability of testing kits and labs, it is quite difficult for every patient with respiratory symptoms to be tested by RT-PCR. Even though in the initial stages of diseases, CXRs are not very sensitive to accurate detection of pulmonary involvement but they can be very helpful in monitoring an increase in lung abnormalities over a period of time [7][8]. CXRs are preferred over CT-scans because the former is more efficient in terms of time and cost involved than the latter. CXR images of COVID-19 patients show specific characteristics which could help in diagnosing and monitoring such patients. Hence, it is widely used in hospitals and medical facilities. Given that the number of patients infected by this disease is far more than the number of radiologists present, there is a need to use an artificial intelligence-based algorithm to aid radiologists in analysing CXR images quickly and predicting COVID-19 severity within a short span of time.

The paper proposes a model which predicts COVID-19 pneumonia severity score. The model involves several steps - data pre-processing, Attention based feature extraction, and severity score prediction using Linear Regression. The first step, data pre-processing involves resizing, normalisation, and NSST [9]. The second step involves use of Attentional Feature Fusion [10] to ResNet-101. In most of the previous work [11][12], features were fused via short skip connection in ResNet which led to very deep network training. Attentional Feature Fusion (AFF) generalises the attention-based feature fusion and helps in providing steady direction, thereby enhancing ResNet performance. The addition operator in the ResNet block is replaced by the AFF module. The third step involves the use of Linear Regression to predict COVID-19 pneumonia severity score using a different set of features. This model can be used to monitor the COVID-19 progress and the response to treatment given to the patients.

The paper is organized as follows. Previous related work has been discussed in Section II below. The materials and the methodology are described in detail in Section III. Section IV provides the experimental results. Finally, in Section V the main conclusions are outlined.

II. RELATED WORKS

In this section, we briefly address the previously proposed approaches that diagnose COVID-19 using medical imaging. These techniques use computed tomography (CT) scans and chest X-ray (CXR) for the initial diagnosis of COVID-19. However, due to better financial viability and portability, CXRs are preferred over CT-scans for the initial imaging modality. For this reason, many healthcare facilities use CXRs as a first diagnostic imaging option for COVID-19 disease monitoring and patient management. Most of the recent works on COVID-19 diagnosis are focused on classification [13], and segmentation [14][15]. [13] developed COVIDX-Net for the classification purpose.

© IEEE 2021. This article is free to access and download, along with rights for full text and data mining, re-use and analysis.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:53:46 UTC from IEEE Xplore. Restrictions apply.

Moreover, deep learning-based models [15] are used to assess the impact of lung segmentation on COVID-19 detection.

However, a more informative visual assessment of CXRs is necessary for COVID-19 diagnosis. Severity assessment of CXR is used as an assistive tool for efficient patient management. It provides the extent of COVID-19 pulmonary abnormalities in the CXRs. The approach of severity analysis can be used for health care management, i.e., escalation or de-escalation of care [16]. Such assessment is done using semiquantitative scoring systems [17][18] that quantitatively estimate the pulmonary involvement to numerical scales. [17]uses scoring systems for CT-scans to quantify the involvement. Yoon et al. [18] performed CXR analysis using a 5-point scale-based chest radiograph scoring system [19] to grade abnormalities. So far, limited work has been done with the use of deep learning approaches in this direction. Recently, Signoroni et al. [20] developed an end-to-end deep learning architecture to predict Brixia score [7]. To obtain the Brixia score, the lung region is subdivided into six zones, and based on the extent of lung abnormalities, a score (range, 0-3) is assigned to each region.

Cohen et al. [21] proposed two COVID-19 severity scores for the assessment: geographic extent score and lung opacity score. Based on the ground glass opacity, a score (range, 0-4; 0 = no lung involvement, 1 $\leq$ 25%, 2 = 25-50%, 3 = 50-75%, 4 $\geq$ 75% lung involvement) is assigned and the sum of score of each lung is calculated which is termed as geometric extent score (range, 0-8). Similarly, lung opacity score (range, 0-6) is the sum of assigned score for each lung depending on the degree of opacity (0 = no opacity; 1 = ground-glass opacity; 2 = consolidation; 3 = white-out). Due to a limited annotated dataset for severity assessment, a CNN model was pre-trained on seven non-COVID-19 CXR datasets consisting of a total of 18 classes to be used as a feature extractor on a small COVID-19 dataset. However, the use of non-COVID-19 datasets and extraneous classes for training significantly affected the performance of the deep learning model.

This work uses Attention Feature Fusion (AFF) based ResNet-101 that is trained using weakly supervised learning to assess the degree of opacity and the extent of lung involvement on posteroanterior (PA) COVID-19 CXR images.

III. Materials and Methods

In this section, we study in detail, the datasets used and the proposed methodology for the severity assessment.

A. Datasets

A COVID-19 image data collection as shown in Fig. 1, was proposed by Cohen et al. [22] consisting of 94 posteroanterior (PA) CXR images (average age; 56±14.8). The images were sourced within the duration of December 2019 to March 2020 and manually annotated by radiologists for Geographic Extent (GE) score and Lung opacity (LO) score. However, due to the limited availability of annotated dataset for this task, we use the pre-training step. The pre-training step avoids overfitting of deep learning model on a very small dataset. For this purpose, we used five publicly available chest X-ray datasets including the Brixia COVID-19 dataset [20](see Table 1). These five datasets were aligned to each other and four classes (COVID-19 pneumonia; Lung Opacity; Consolidation; Infiltration) from the datasets were used for the pre-training step.

TABLE I. Different public datasets used in the experimentation

Dataset Used	Classes Used
Brixia COVID-19 [20]	COVID-19 pneumonia
CheXpert - Stanford University [23]	Lung Opacity, Consolidation
MIMIC-CXR – MIT [24]	Lung Opacity, Consolidation
OpenI [25]	Lung Opacity
ChestX-ray8 - National Institutes of Health (NIH) [26]	Infiltration

B. Severity Assessment

We introduced a multi-stage framework that assesses the severity of the infection in the CXR region of COVID-19 patients. The framework is designed to learn and perform different tasks such as: data pre-processing, feature extraction, and the severity score prediction: geography extent score and lung opacity score. The proposed methodology is discussed in detail below.

a. Data Pre-processing

Most of the research conducted on the CXR use entire raw data for the training purpose. However, there is a major drawback of using untreated data: the existence of irrelevant objects including such as symbols, bright texts, varying spatial resolutions in the CXR images. This adds to the unnecessary noise that limits the performance of the deep learning networks. Below are some data pre-processing techniques to alleviate some of the aforementioned issues.

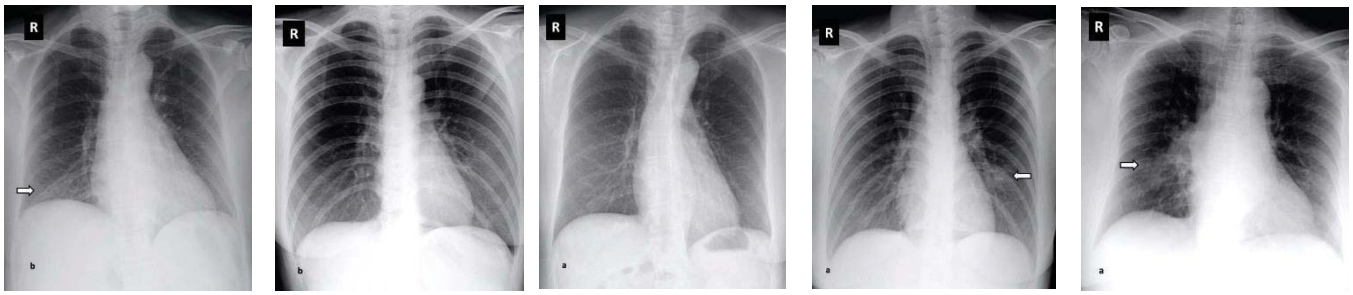


Fig. 1. The illustration of some example chest X-rays from the posteroanterior (PA) COVID-19 CXR dataset [32]

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:53:46 UTC from IEEE Xplore. Restrictions apply.

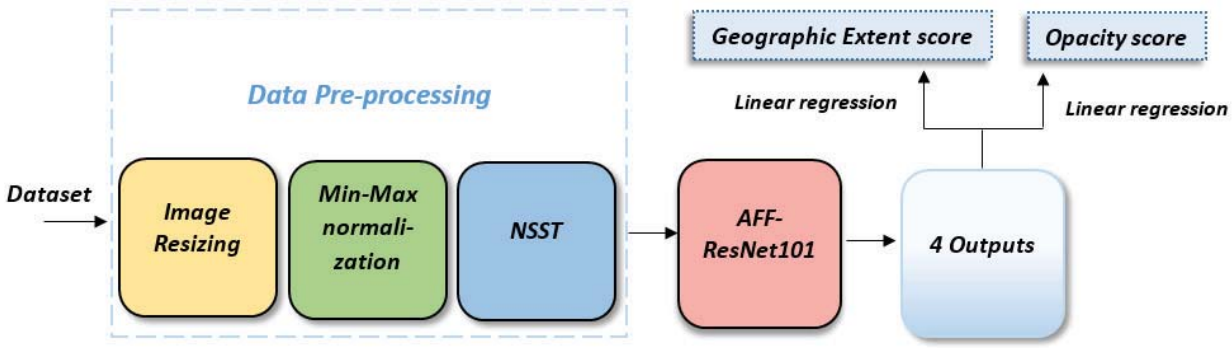


Fig. 2. Block diagram of methodology for severity assessment

- **Image resizing** and **min-max normalization** is applied on X-ray images to ensure uniformity across different datasets.
- **Non-subsample shearlet transform** (NSST) [9] consists of Laplacian pyramid filters and shift-invariant filters that are the extension of the wavelet transform. It decomposes the noisy X-ray image into multi-scale sub-band and multi-directional images. Further, these images are de-noised with the help of filters and combined to get an output X-ray image using an inverse shearlet transform.

b. Attention-based Feature Extraction

Here, we apply Attentional Feature Fusion (AFF) [10], a multi-scale attention-based mechanism to ResNet-101 [12]. The feature fusion operator i.e., the addition operator in ResNet block (ResBlock) is replaced with the attention mechanism. The module addresses the issue of contextual aggregation and enhances contextual awareness by feature-wise modulation and refinement. As it is challenging to train a CNN model on small datasets, we adopt two-stage training to address the issue: Initially, ResNet-101 with Attentional Feature Fusion (AFF) module is trained on the six datasets as mentioned in Table 1. Further we perform supervised fine-tuning on the smaller dataset using the initially trained AFF-ResNet101.

IV. Experiments

After applying the aforementioned pre-processing steps on the five chest X-ray datasets, weakly-supervised pretraining was performed using Attentional Feature Fusion (AFF) based ResNet-101 and further, transfer learning is applied to 94 posteroanterior (PA) CXR images of COVID-19 patients. In the experimentation, five-fold cross-validation is adopted to train the deep learning model with RMSprop optimizer and the initial learning rate is chosen at 0.0001. The implementation of the proposed technique is done using the Pytorch framework on two Nvidia RTX 2080Ti GPU. As per evaluation strategies used in the evaluation of previous literature work, linear regression is trained on the different sets of features to predict the lung involvement and the degree of opacity. For the purpose, the training sample of 94 posteroanterior (PA) CXR images is divided into training set (50%) and test set (50%). Further, performance metrics including R Square (R^2), Pearson correlation, and mean squared error (MSE) are adopted for the model evaluation. Block diagram representation of the methodology is shown in Fig. 2. The comparison between model's prediction and the ground truth for geographic extent and opacity (as illustrated in Fig. 3) shows that the most of the points lie near

TABLE II. Comparison of the different variants of our weakly supervised approach with other methodologies on Posteroanterior (PA) CXR images

	Task	Opacity Score			Geographical Extent		
Method		"lung opacity" output	4 outputs	Intermediate network features	"lung opacity" output	4 outputs	Intermediate network features
Cohen et al. [21]	Pearson Correlation	0.78±0.04	0.78±0.04	0.66±0.08	0.80±0.05	0.79±0.05	0.74±0.08
	R square	0.58±0.09	0.58±0.09	0.25±0.21	0.60±0.09	0.57±0.10	0.43±0.16
	MSE	0.86±0.11	0.87±0.12	1.54±0.28	2.06±0.34	2.17±0.37	2.88±0.58
ResNet-101 + Regression (Ours)	Pearson Correlation	0.79±0.04	0.79±0.04	0.67±0.09	0.81±0.06	0.80±0.05	0.75±0.09
	R square	0.61±0.09	0.58±0.09	0.26±0.21	0.63±0.09	0.57±0.11	0.46±0.15
	MSE	0.88±0.12	0.86±0.11	1.56±0.27	2.09±0.37	2.19±0.36	2.90±0.57
AFF-ResNet-101 + Regression (Ours)	Pearson Correlation	0.85±0.05	0.86±0.05	0.71±0.10	0.86±0.05	0.86±0.05	0.71±0.10
	R square	0.66±0.11	0.64±0.10	0.32±0.23	0.66±0.10	0.62±0.10	0.50±0.17
	MSE	0.79±0.09	0.81±0.12	1.47±0.29	1.98±0.32	2.06±0.35	2.69±0.59

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:53:46 UTC from IEEE Xplore. Restrictions apply.

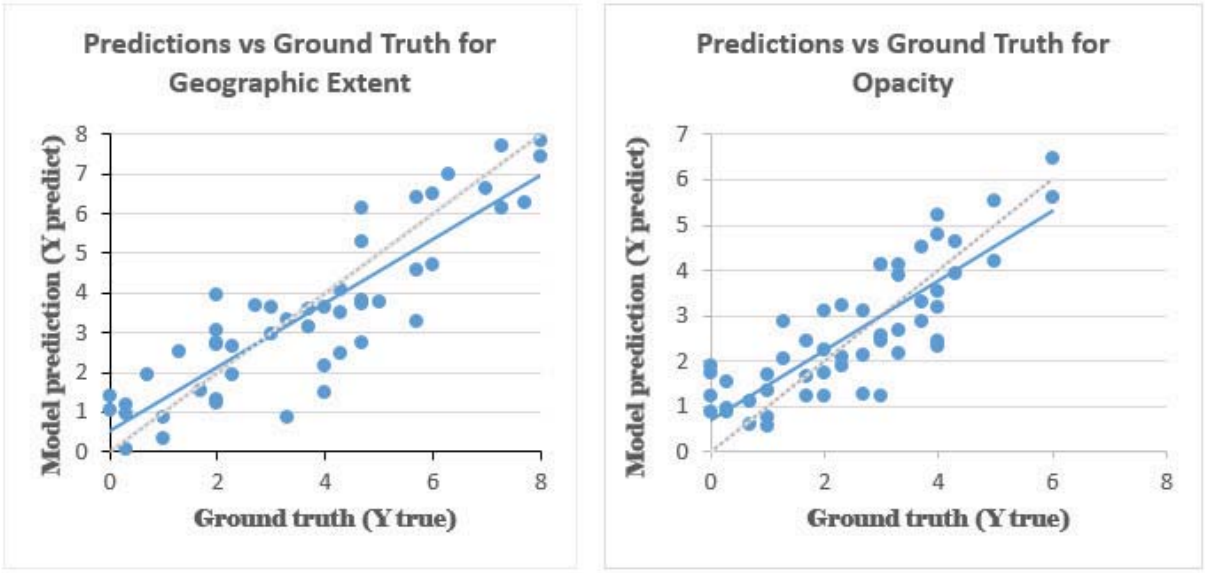


Fig. 3. Two scattered plots representing the score prediction made by model vs the ground truth (manual annotation by radiologist) for geographic extent and opacity respectively. In this graph, blue line represents linear regression line while the dotted grey line represents the exact prediction.

the grey line representing perfect prediction.

As observed in Table 2, AFF-ResNet-101 is able to achieve improved performance over the ResNet-101 based weakly-supervised approach and Cohen et al. [1] that offers no such attention mechanism. It is attributed to the fact that Attention Feature Fusion (AFF) module replaces the conventional addition operation. It enhances the feature fusion in skip connections of ResNet-101 and boosts the model's ability to identify regions of utmost importance. Moreover, Cohen et al. [21] trained DenseNet-101 on a large set of non-COVID-19 datasets that were aligned for 18 classes. The use of non-COVID-19 datasets and irrelevant classes create unnecessary confusion in the feature extraction process. However, our approach enables the initial and target dataset to be similar and relevant for the first round of training i.e., the training on five datasets. It provides an unbiased approach for an efficient feature extractor that further, allows better learning for the linear regression model.

V. Conclusions

The present paper proposes a pipeline that uses Attention-based weakly supervised learning for automatic severity assessment based on the degree of opacity and geographical extent. To the best of our knowledge, this is the first study that uses an attention mechanism to gauge the level of severity of infection in COVID-19 positive CXR. Moreover, our framework was able to outperform the existing state-of-the-art method on the posteroanterior (PA) COVID-19 CXR dataset. In the future, our aim would be to extend the scope of the proposed system to the other pulmonary diseases and investigate further improvements for better efficiency and robustness.

References

[1] The Lancet Respiratory Medicine, "COVID-19 transmission—up in the air," Lancet Respir. Med., vol. 8, no. 12, p. 1159, Dec. 2020, doi: 10.1016/S2213-2600(20)30514-2.

[2] K. Chaturvedi, D. K. Vishwakarma, and N. Singh, "COVID-19 and its impact on education, social life and mental health of students: A survey," Child. Youth Serv. Rev., vol. 121, p. 105866, Feb. 2021, doi: 10.1016/j.childyouth.2020.105866.

[3] [C. Huang et al., "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," Lancet, vol. 395, no. 10223, pp. 497–506, Feb. 2020, doi: 10.1016/S0140-6736(20)30183-5.

[4] N. Zhu et al., "A Novel Coronavirus from Patients with Pneumonia in China, 2019," N. Engl. J. Med., vol. 382, no. 8, pp. 727–733, Feb. 2020, doi: 10.1056/NEJMoa2001017.

[5] E. Mahase, "Coronavirus: covid-19 has killed more people than SARS and MERS combined, despite lower case fatality rate," BMJ, p. m641, Feb. 2020, doi: 10.1136/bmj.m641.

[6] D. Wang et al., "Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan, China," JAMA, vol. 323, no. 11, p. 1061, Mar. 2020, doi: 10.1001/jama.2020.1585.

[7] A. Borghesi and R. Maroldi, "COVID-19 outbreak in Italy: experimental chest X-ray scoring system for quantifying and monitoring disease progression," Radiol. Med., vol. 125, no. 5, pp. 509–513, May 2020, doi: 10.1007/s11547-020-01200-3.

[8] R. Yasin and W. Gouda, "Chest X-ray findings monitoring COVID-19 disease course and severity," Egypt. J. Radiol. Nucl. Med., vol. 51, no. 1, p. 193, Dec. 2020, doi: 10.1186/s43055-020-00296-x.

[9] A. Zhang, H. Jiang, L. Ma, Y. Liu, and X. Yang, "A Shearlet-based algorithm for quantum noise removal in low-dose CT images," Mar. 2016, p. 97843O, doi: 10.1117/12.2216562.

[10] Y. Dai, F. Gieseke, S. Oehmcke, Y. Wu, and K. Barnard, "Attentional Feature Fusion," Sep. 2020, [Online]. Available: http://arxiv.org/abs/2009.14082.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Identity Mappings in Deep Residual Networks," 2016, pp. 630–645.

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[13] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images," Mar. 2020, [Online]. Available: http://arxiv.org/abs/2003.11055.

[14] A. Degerli et al., "COVID-19 Infection Map Generation and Detection from Chest X-Ray Images," Sep. 2020, [Online]. Available: http://arxiv.org/abs/2009.12698.

[15] L. O. Teixeira et al., "Impact of lung segmentation on the diagnosis and explanation of COVID-19 in chest X-ray images," Sep. 2020, [Online]. Available: http://arxiv.org/abs/2009.09780.

[16] J. M. Courtney and A. Bax, "Hydrating the Respiratory Tract: An Alternative Explanation Why Masks Lower Severity of COVID-19 Disease.," medRxiv Prepr. Serv. Heal. Sci., Dec. 2020, doi: 10.1101/2020.12.23.20248671.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:53:46 UTC from IEEE Xplore. Restrictions apply.

[17] F. Pan et al., “Time Course of Lung Changes at Chest CT during Recovery from Coronavirus Disease 2019 (COVID-19),” Radiology, vol. 295, no. 3, pp. 715–721, Jun. 2020, doi: 10.1148/radiol.2020200370.

[18] J. Zhu, B. Shen, A. Abbasi, M. Hoshmand-Kochi, H. Li, and T. Q. Duong, “Deep transfer learning artificial intelligence accurately stages COVID-19 lung disease severity on portable chest radiographs,” PLoS One, vol. 15, no. 7, p. e0236621, Jul. 2020, doi: 10.1371/journal.pone.0236621.

[19] E. Taylor et al., “A chest radiograph scoring system in patients with severe acute respiratory infection: a validation study,” BMC Med. Imaging, vol. 15, no. 1, p. 61, Dec. 2015, doi: 10.1186/s12880-015-0103-y.

[20] A. Signoroni et al., “End-to-end learning for semiquantitative rating of COVID-19 severity on Chest X-rays,” Jun. 2020, [Online]. Available: http://arxiv.org/abs/2006.04603.

[21] J. P. Cohen et al., “Predicting COVID-19 Pneumonia Severity on Chest X-ray With Deep Learning.,” Cureus, vol. 12, no. 7, p. e9448, Jul. 2020, doi: 10.7759/cureus.9448.

[22] J. P. Cohen, P. Morrison, and L. Dao, “COVID-19 Image Data Collection,” Mar. 2020, [Online]. Available: http://arxiv.org/abs/2003.11597.

[23] J. Irvin et al., “CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison,” Jan. 2019, [Online]. Available: http://arxiv.org/abs/1901.07031.

[24] A. E. W. Johnson et al., “MIMIC-CXR-JPG, a large publicly available database of labeled chest radiographs,” Jan. 2019, [Online]. Available: http://arxiv.org/abs/1901.07042.

[25] D. Demner-Fushman et al., “Preparing a collection of radiology examinations for distribution and retrieval.,” J. Am. Med. Inform. Assoc., vol. 23, no. 2, pp. 304–10, Mar. 2016, doi: 10.1093/jamia/ocv080.

[26] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, “ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases,” in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jul. 2017, pp. 3462–3471, doi: 10.1109/CVPR.2017.369.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:53:46 UTC from IEEE Xplore. Restrictions apply.

2021 2nd International Conference for Emerging Technology (INCET)
Belgaum, India. May 21-23, 2021

Design and Study of Dadda Multiplier by using 4:2 Compressors and Parallel Prefix Adders for VLSI Circuit Designs

Ayush Kumar
Department of ECE
Delhi Technological University
New Delhi, India
ayushkumardtu@gmail.com

Ankit Kumar
Department of ECE
Delhi Technological University
New Delhi, India
akankitkmr53@gmail.com

Dr. Deva Nand
Department of ECE
Delhi Technological University
New Delhi, India
devkamboj@dce.ac.in

2021 2nd International Conference for Emerging Technology (INCET) | 978-1-7281-7029-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/INCET51464.2021.9456283

***Abstract*—In this paper, a new Dadda Multiplier is designed using 4:2 compressors and parallel prefix adders. From study and analysis in the literature reported in the recent past it is well known that a significant functional component in any processing module is a multiplier. In the processor, the multiplier takes up the majority of the hardware consumed as compared to any other arithmetic operations. Among the existing multiplication methods, Dadda multiplication method is advantageous in terms of reduction in delay.**

The main aim of a quality multiplier is to be as fast and consume as low area as possible. To decrease the delay of Dadda multiplier we have employed 4:2 compressor in the reduction stage. In the proposed system, parallel prefix adders are used to add the final stage of the partial product. In this paper, 4 Dadda multipliers employing 4:2 compressor block are proposed using Sklansky adder (SA), Kogge-Stone adder (KSA), Brent-Kung adder (BKA), and Ladner-Fischer adder (LFA). The proposed multiplier designs and conventional design are simulated on Xilinx Vivado 2018.2. The proposed designs are analyzed concerning conventional multiplier structure in terms of parameters delay (ns) and area (No. of LUTs).

Keywords— Sklansky Adder, Dadda, Kogge-Stone Adder, Parallel Prefix Adder, Brent-Kung Adder, Ladner-Fischer Adder

I. INTRODUCTION

In DSP (Digital Signal Processing) various essential functions are performed in which multiplication is the most important one from all of the other fundamental functions like arithmetic addition and subtraction, these operations can be performed much easily as compared to the requirements for multiplication. It has been observed that the delay in DSP (Digital Signal Processing) system majorly depends on the delay time of the multiplier circuit. The demand has been increasing for High-speed processors as a result of expanding computers and signal processing applications causes a need for high-speed multiplier.

In designing a multiplier, low power consumption is also a crucial issue, so in order to reduce significant power consumption, it is better to minimize the number of operations thereby decreasing dynamic power which is an important part of total power consumption so the demand for low power and high-speed multiplier has increased.

At present, technology is moving along very quickly in a small period of time. The circuits which are manufactured consist of large number of components with high speed, low area, and also low power consumption. Hence, In the design of a circuit, area, speed, and power plays a crucial role [1], [2]. Which leads to the demand of a multiplier circuit which meets these current trends.

In the literature survey of multiplier design, we have gone through different multiplication algorithms like Array multiplier, Binary multiplier, Wallace Tree multiplier, Dadda multiplier.[3] Dadda multiplier algorithm is dominant in different types of multipliers [4]. Dadda multiplier works as same as Wallace Tree Multiplier [5] but it has been observed that Dadda Multiplier is a bit faster in nature. Another advantage of Dadda Multiplier is that the number of components are less as compared to Wallace tree structure.[4]

Prefix adders are derived from parallel prefix circuit concept which comes up with a strong theoretical basis for a broad range of design trade-offs between area, delay and writing complexity.[9]

In the proposed design of Dadda multiplier, the operation of the multiplier is enhanced by employing the 4:2 compressor, due to which the number of levels in partial product reduction stage gets minimized.

In this paper, a different design of Dadda multiplier is proposed in which parallel prefix adders are employed to calculate the final two rows of partial products which are made after reduction from reduction tree stage with the help of 4:2 compressors, full adders and half adders to produce final product terms. All the PPAs are based on carry look-ahead adder theory of generating and propagating carry bits.[10] Kogge Stone adder[11], Sklansky adder[12], Brent-Kung adder[13], Ladner-Fischer adder [14]are some of the adders utilized in the design of Dadda multiplier.

This paper is arranged into various sections. Section I is the Introduction. Section II throws light on the comprehensive functionality of the Dadda multiplier and gives brief information about 4:2 compressor. Section III contains numerous subdivisions which express the in-depth functionality of different parallel prefix adders. In section IV the outcome analysis is distinctly described. Section V consists of the conclusion and future scope of the proposed multipliers in this paper.

II. DADDA MULTIPLIER

Dadda algorithm was invented in the year 1965 by Computer scientist Luigi Dadda.[4] Dadda Algorithm strives

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:00:04 UTC from IEEE Xplore. Restrictions apply.

to reduce the propagation delay and numbers of basic logic gates utilized, unlike Wallace algorithm which only tries to reduces delay as much as possible on each layer. Thus Dadda Multiplier is more efficient.

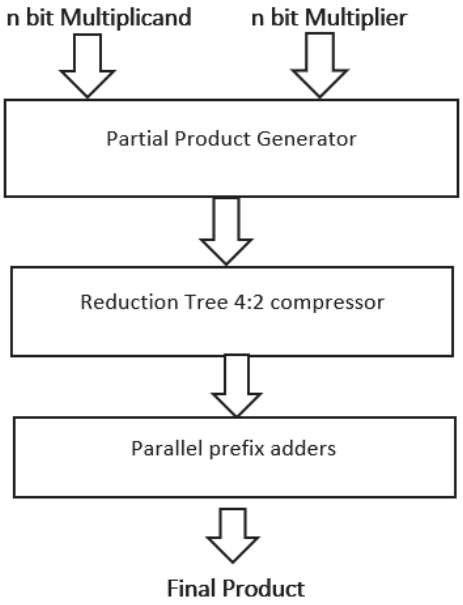


Fig. 1. Block Diagram for Dadda Algorithm

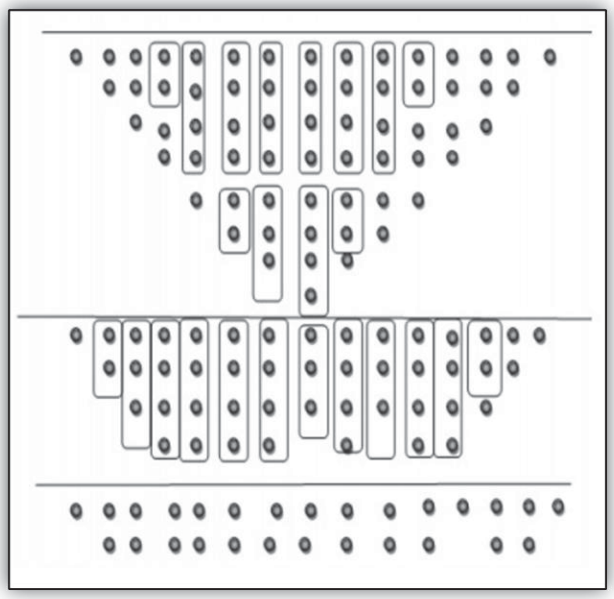


Fig. 2. Dadda Algorithm Dot Diagram[7]

Steps to multiply 2 numbers using Dadda algorithm:

1. Form the partial product by multiplying(and-ing) the multiplicand and multiplier bits by bits, resulting in a matrix known as partial product matrix.
2. The multiplication products of the first stage carry different weights which depends on the position of the multiplied bits.
3. Employing full adders and half adders and 4:2 compressor reduces the number of stages of partial product reduction until two rows are left.

Group the products into 2 numbers and perform addition using parallel prefix adder.

A. 4:2 Compressor:

In multipliers for adding partial product terms, compressors are widely used so as to lessen the number of operands. There are 5 inputs in 4:2 compressor, among which one input is the carry generated in the previous stage and it has 3 outputs which are sum, cout and carry bits.[6] In multiplier when we add many partial products then to reduce the number of operations it is convenient to use 4:2 compressors instead of 3:2 compressors.[8]

Cout = (S1andS2)or(S2andS3)or(S3andS1)

= ((S1xorS2)andS3)or(S1andS2)

Sum = S1xorS2xorS3xorS4xorCin

Carry = (S4xorCin)and(S1xorS2xorS3)or(S4andCin)

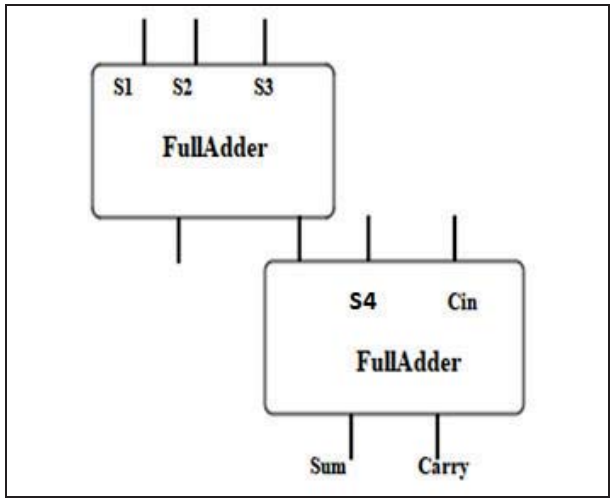


Fig. 3. Traditional Structure of 4:2 Compressor[7]

III. PARALLEL ADDERS

A. Parallel Prefix adder:

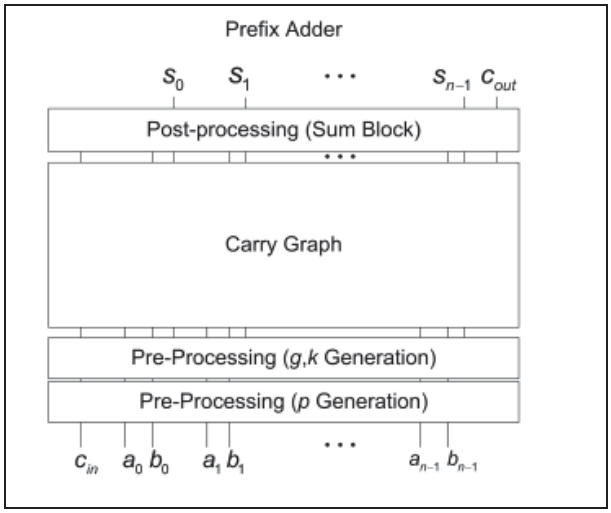


Fig. 4. Prefix Adder[9]

In this paper the proposed Dadda multiplier configuration is improved by employing parallel prefix adders to generate

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:00:04 UTC from IEEE Xplore. Restrictions apply.

the sum of the last two partial products in the final stage addition process to obtain the final product. The main variance between the parallel adders and prefix adders is that the sum and carry generation is performed in the corresponding one-bit block in parallel adder while on the other hand in prefix adder, sum and carry generation is considered as a whole in the carry graph.[9] The carry graph uses the prefix circuit and this lead to the creation of the name "Prefix Adder".

In the pre-processing stage of the parallel prefix adder, the carry generation and carry propagation signals are calculated for the carry graph by each pair of input bits. The generate and propagate equations are mentioned below:

Generate: g = (x and y)

Propagate: p = (x xor y)

After the computation, carries are divided into small packages. These are known as processing components and there following equations are mentioned below:

Pi:k = (Pi:j and Pj-1:k)

Gi:k = (Gi:j or Gj-1:k and Pi:j)

In the post-processing block, Sum and final output carry are calculated:

Si = (Pi xor Ci-1)

Carry graph is important in parallel prefix adders, and it has 3 different components which are mentioned below:

1. Black Cell

2. Gray Cell

3. Buffer

Black cells calculate both carry generate and propagate bits.[15] Grey cells calculate carry generate a signal only. Buffers are used to maintain the loading effect.[15]

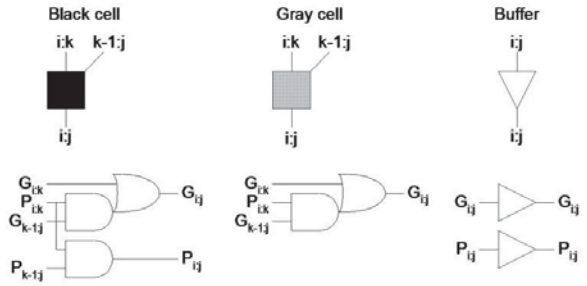


Fig. 5. Carry generation and propagation structures in Parallel Prefix Adders

1) Kogge-Stone Adder:

Kogge stone adder is originated from the logic of carry look ahead adder. It performs a very fast logical addition operation and can be used for arithmetic circuits with top performance. The prime benefit of Kogge Stone adder is it has minimum fanouts. the limitations of this adder are that the operational power is more and it occupies a larger area. The structure of Kogge Stone adder for 16-bits inputs consists of 8 buffers, 15 grey cells and 34 black cells.

2) Sklansky Adder:

The Sklansky adder is based on divide and conquer tree concept. It has a simple prefix structure. In the structure of carry prefix network there is a trade-off between fan out and logic depth, it has high fan out and minimum logic depth. The structure of Sklansky adder for 16-bits inputs consists of 20 buffers, 15 grey cells and 17 black cells.

3) Brent-Kung Adder:

Brent-Kung adder has logic depth in maximum and the number of nodes in minimum. This will also reduce the delay when implemented higher bits adders without compromising the power performance of the adder. It has a higher degree of regularity and has very less wiring crowding which leads to improved performance and requires less chip area to design as compared to Kogge Stone adder. The structure of Brent-Kung adder for 16-bits inputs consists of 20 buffers, 15 grey cells and 12 black cells.

4) Ladner-Fischer Adder:

By hybrid of Sklansky and Brent Kung PPAs, Ladner-Fischer adder is made. This adder design has a large fan-out and has minimum logic depth. The structure of Ladner-Fischer adder for 16-bits inputs consists of 19 buffers, 15 grey cells and 12 black cells.

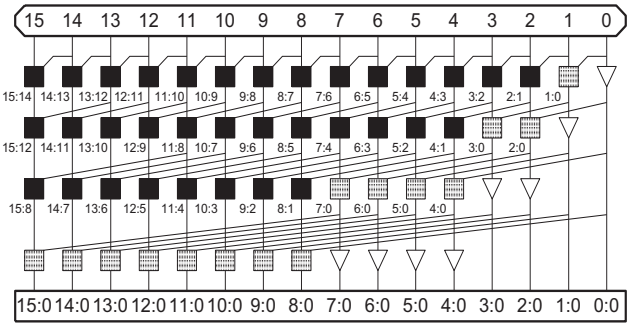


Fig. 6. Kogge-Stone Adder (KSA)

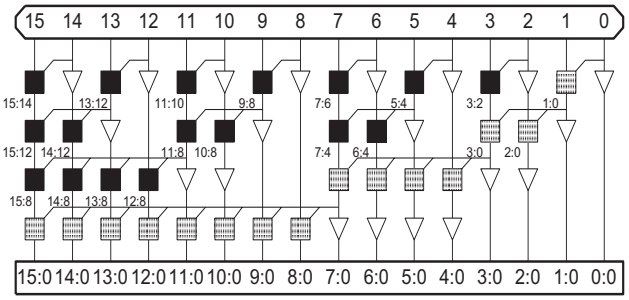


Fig. 7. Sklansky Adder (SA)

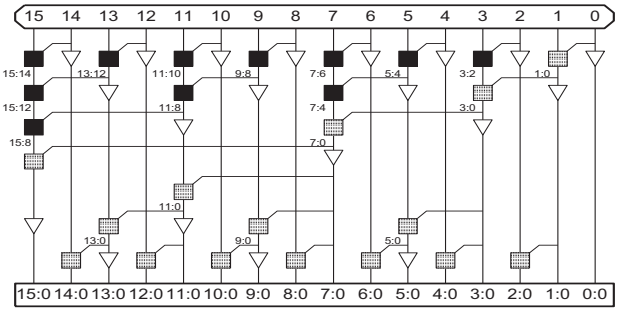


Fig. 8. Brent-Kung Adder (BKA)

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:00:04 UTC from IEEE Xplore. Restrictions apply.

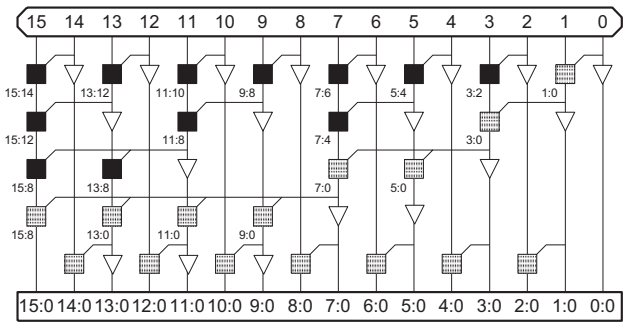


Fig. 9. Ladner-Fischer Adder (LFA)

IV. SIMULATED AND SYNTHESIZED RESULT ANALYSIS

Four Dadda multiplier employing 4:2 compressor structures utilizing four different PPA's (Sklansky Adder, Kogge-Stone adder, Brent-Kung adder and Ladner Fischer adder) are designed for the size of input of 16-bits each in this paper using Verilog HDL. All the multipliers simulation waveforms are shown in Fig10-Fig14. The four proposed multipliers synthesis report results are compared with the synthesis report of Conventional Dadda multiplier using Carry Look Ahead Adder (CLA) in terms of parameter delay in nanoseconds and area (number of LUTs). All the multipliers designed are synthesized and simulated using Xilinx Vivado 2018.2 and selecting Kintex Ultrascale+ xcku15p-ffve1760-1LV-i.

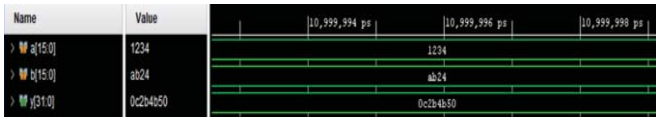


Fig. 10. Waveform simulation of 16-bit Conventional Dadda multiplier employing 4:2 compressor.

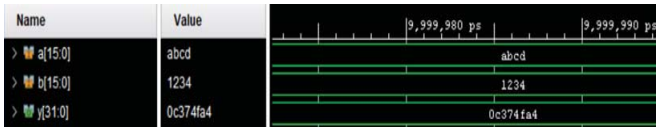


Fig. 11. Waveform simulation of 16-bit Dadda multiplier employing 4:2 compressor and Kogge-Stone Adder.

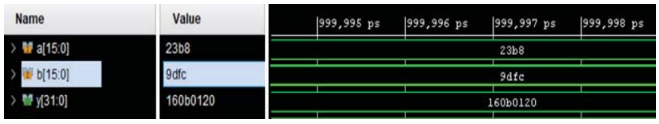


Fig. 12. Waveform simulation of 16-bit Dadda multiplier employing 4:2 compressor and Sklansky Adder.

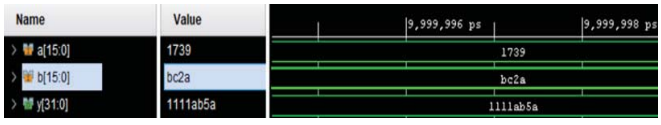


Fig. 13. Waveform simulation of 16-bit Dadda multiplier employing 4:2 compressor and Brent-Kung Adder.

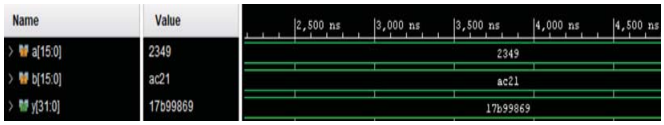


Fig. 14. Waveform simulation of 16-bit Dadda multiplier using 4:2 compressor and Ladner-Fischer Adder.

TABLE I. DELAY(NS) AND AREA (NUMBER OF LUTS) OF DIFFERENT TYPE OF DADDA MULTIPLIER STRUCTURES EMPLOYING 4:2 COMPRESSOR AND DIFFERENT PPAS

Input Size	Type of Multiplier	Delay(ns)	Area (No. of LUT's)
16-bit	Conventional Dadda Multiplier employing 4:2 compressor	9.944	450
	Dadda multiplier employing 4:2 compressor and Kogge Stone adder	8.879	375
	Dadda multiplier employing 4:2 compressor and Sklansky adder	9.207	373
	Dadda multiplier employing 4:2 compressor and Brent Kung adder	9.214	365
	Dadda multiplier employing 4:2 compressor and Ladner Fischer adder	9.201	371

In table 1, a total of five multipliers i.e., four type of proposed multipliers using PPAs and one conventional multiplier utilizing CLA are synthesized and simulated. The synthesis report of each multiplier consists of area details in terms of parameter delay in nanoseconds and the number of LUTs occupied. From table 1, it can be inferred that:

1. Dadda multiplier employing 4:2 compressor and Kogge Stone adder has the least delay, but in terms of area it occupies a greater number of LUTs when compared to other proposed multipliers.
2. The traditional multiplier employing CLA has the maximum delay and number of LUTs occupied.
3. The Dadda multiplier employing 4:2 compressor and Brent Kung adder has the least number of LUTs used but its delay is more than other types of multipliers with PPA's but is less when compared with the traditional structure.
4. The Dadda multiplier employing 4:2 compressor and Ladner Fischer adder lies in between all the other adders compared as its delay and number of LUT's is moderate.

V. CONCLUSION AND FUTURE SCOPE

In this paper, a 16x16 Dadda multiplier employing 4:2 compressor and different PPA's has been realized using Verilog with the help of Xilinx Vivado. The 4:2 compressor improves the performance of the Dadda multiplier during the partial product reduction stage while PPA's help in minimizing the delay in the final product stage. The four-multipliers structure designed in this paper of Dadda multiplier using 4:2 compressor and 4 different PPA's is compared with Conventional Dadda multiplier in terms of parameters like delay and area requirement in term of LUT's. There is a significant improvement in terms of the Delay parameter and number of LUT's used which is observed from table 1. In the multipliers proposed in this paper, Dadda multiplier employing 4:2 compressor and Kogge Stone adder has the least delay i.e., it's the fastest. But the number of LUT's used by it is more as compared to other multiplier structures. While for the area parameter, Dadda multiplier using 4:2 compressor and Brent Kung adder has the least area in terms of the number of LUT's used amongst the other proposed structures.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:00:04 UTC from IEEE Xplore. Restrictions apply.

So, the Dadda multiplier employing 4:2 compressor and Kogge Stone adder circuit can be used for multiplier application which needs to be fast but can cost a little extra area. Whereas for area constricted application Dadda multiplier employing 4:2 compressor and Brent Kung adder can be a good alternative. Finally, if a balanced multiplier application i.e., which needs to be moderately fast but at the same time isn't at the expensive side can use Dadda multiplier using 4:2 compressor and Ladner Fischer adder. In the future scope of this work, compressors like 5:2, 7:2, etc. can be used or the size of input can be increased (i.e., 32-bits, 64-bits).

REFERENCES

[1] N. Weste and D. Harris, CMOS VLSI Design. Reading, MA: Addison Wesley, 2004.

[2] N. Weste and K. Eshragian, Principles of CMOS VLSI Designs: A System Perspective, 2nd ed., Addison-Wesley, 1985-1993.

[3] J. M. Rabaey, Digtal Integrated Circuits—A Design Perspective. Upper Saddle River, NJ Prentice-Hall, 2001.

[4] L. Dadda, "Some schemes for parallel multipliers," Alta Frequenza, Vol. 34, pp. 349-356, May, 1965.

[5] C. S. Wallace, "A suggestion for a fast multiplier," IEEE Trans. Electronic Computers, Vol. 13, pp. 14-17, Feb., 1964.

[6] U. Penchalaiah and S. K. VG, "Design of High-Speed and Energy-Efficient Parallel Prefix Kogge Stone Adder," 2018 IEEE International Conference on System, Computation, Automation and Networking (ICSCA), Pondicherry, 2018, pp. 1-7.

[7] M. V. S. Nandam and S. Alluri, "High Performance 32-bit Dadda Multiplier Using EDA," 2020 IEEE 7th International Conference on Smart Structures and Systems (ICSSS), Chennai, India, 2020, pp. 1-5, doi: 10.1109/ICSSS49621.2020.9201958.

[8] Rupendra Singh and Neeta Pandey, “A Low Power and Area Efficient Design of Dadda Multiplier by exploring 4:2 Compressors and Brent-Kung Adder,” International Journal of Advanced Science and Technology Vol. 29, No. 3, (2020), pp. 10044 – 10054

[9] Choi, Youngmoon & Jr, Earl. (2005). Parallel Prefix Adder Design with Matrix Representation.90-98. 10.1109/ARITH.2005.35. 17th IEEE Symposium on Computer Arithmetic (ARITH-17 2005), 27-29 June 2005, Cape Cod, MA, USA

[10] Sarnala Butchibabu, Sannikanti Kishore Babu, 2014, Design and Implementation of Efficient Parallel Prefix Adders on FPGA, International Journal of Engineering Research & Technology (IJERT) Volume 03, Issue 07 (July 2014),

[11] P. M. Kogge and H. S. Stone, “A Parallel Algorithm for the Efficient Solution of a General Class of Recurrence Equations,” IEEE Trans. Computers, Vol.C-22, 1973, pp.786 793.

[12] J. Sklansky, “Conditional-sum addition logic,” IRE Transactions on Electronic Computers, vol. 9, pp. 226–231, 1960.

[13] R. P. Brent and H. T. Kung, “A Regular Layout for Parallel Adders,” IEEE Trans. Computers, Vol. C-31, 1982, pp. 260-264.

[14] R.E. Ladner and M.J. Fischer, “Parallel Prefix Computation,” JACM, Vol.27-4, 1980, pp. 831-838.

[15] N Aruna Kumari, M Sai Srinivas, K Aravind, “Implementatio n of 64 Bit Arithmetic Adders,” International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7, Issue-5S4, February 2019.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:00:04 UTC from IEEE Xplore. Restrictions apply.

2021 2nd International Conference for Emerging Technology (INCET)
Belgaum, India. May 21-23, 2021

Design Of High Bandwidth Circularly Polarised Antipodal Vivaldi Array for 5G Applications

Anirudh Nakra
Dept. of Electronics and Communication
Delhi Technological University
New Delhi, India
anirudhnakra4@gmail.com

Abhijeet Vats
Dept. of Electronics and Communication
Delhi Technological University
New Delhi, India
abhijeetvats01@gmail.com

Asok De
Professor,
Dept. of Electronics and Communication
Delhi Technological University
New Delhi, India
asok.de@gmail.com

2021 2nd International Conference for Emerging Technology (INCET) | 978-1-7281-7029-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/INCET51464.2021.9456438

***Abstract*— In this paper, a circularly polarized antenna array is proposed that has high gain and excellent axial ratio properties. The array is composed of multiple UWB antipodal slotted Vivaldi antennas that have been designed for a resonant frequency of 8.5 GHz and using orthogonal excitations to the array, circular polarization has been successfully achieved. Modified Wilkinson power splitter has been utilized to work over the range of 4-12 GHz and phase shifting has been achieved using aperture coupling techniques. The S parameters for the constructed system are excellent with return loss as low as -25 dB and an axial ratio of almost 0.098 dB at the resonant frequency.**

Keywords—Circular Polarization, Antipodal Vivaldi Antenna, Wilkinson divider, Axial ratio

I. INTRODUCTION

Circularly polarized antennas provide many advantages over a classical arrangement. For one, it allows a greater possibility of a good link between the transmitter and the receiver. It allows good LoS linkages and therefore is an important way forward for the antenna world. Recently with the proliferation of UWB circularly polarized antennas, there has been a rise in research in such areas. Good outputs and axial ratios can be observed over large bandwidths and this has made the demand for an antenna array that much more pivotal.

Some great research has been done on integrating aperture coupling techniques with UWB antennas to create antennas that work better than those using delay line circuits. Narbudowicz et all have employed a fusion system of good isolation power splitters with UWB phase shifters and antipodal Vivaldi antennas.

5G and further generations will benefit from CP antennas due to their extreme resistance to reflection, absorption and other interferences and mismatched polarization. In the paper, we have used an antipodal Vivaldi antenna with slots to maximize the bandwidth while improving isolation between the various

Components using a six stage power divider. We have also compared various different planar antennas such as a rectangular microstrip antenna and Quasi-Yagi antenna before finally landing on the antipodal Vivaldi antenna due to its better bandwidth and antenna efficiency results.

II. DESIGN

We chose the antipodal Vivaldi antenna as our working piece due to its high bandwidth. We achieved a gain of 7.6 dB with an antenna efficiency of 98.7% while also having a practical bandwidth of 7.8 GHz. This all round excellence and relative ease of fabrication made it an obvious choice over the Quasi Yagi antenna which had a slightly better efficiency but compromised on the bandwidth. Furthermore, research on antipodal Vivaldi has come to a point where its characteristics are quite suitable for 5G applications and creating a circularly polarized system would be beneficial for future work.

Two power splitters were tested in the feed network: a 3 stage Wilkinson power splitter and a 6 stage Wilkinson power splitter. The 6 stage Wilkinson power divider designed on a substrate of thickness 0.508 mm and dielectric constant 2.2 and it showed better isolation between the output ports compared to the 3 stage divider.

A Multi Sectional structure with different impedances were used to achieve the operation of the power divider in the required frequency range. The values of the resistances were calculated by the formulas given in [6]. Using the LineCalc tool provided in the schematic interface of Keysight ADS, we were able to define our control parameters to calculate the optimal length and width of the different stages. After calculating the characteristic impedances of the various stages mathematically and using TX line, we calculated, and fine-tuned the dimensions to those that gave us the best results. These widths and lengths were then translated into a 3d model in CST design studio using a delta separation of 1.3mm between the two isolated arms. By analyzing the S parameter curves, we saw that we were able to achieve an isolation greater than 28 dB which was better than the isolation from the splitters in both [3] and [4], two studies we chose as our standard.

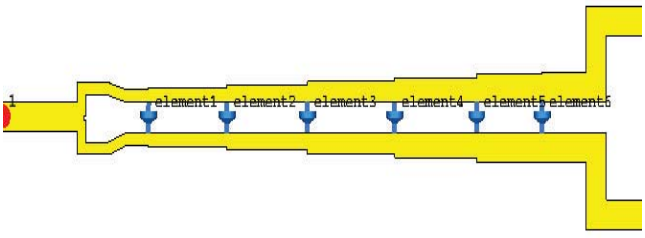


Fig. 1. Wilkinson Power divider

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:00:44 UTC from IEEE Xplore. Restrictions apply.

TABLE I. WILKINSON POWER SPLITTER DIMENSIONAL PARAMETERS AND RESISTOR VALUES

Resistors	Dimensional Parameters of Wilkinson power divider		
	Resistor values	*Width(mm)*	*length*
R1	153	W1=0.42	L1=5.20
R2	166	W2=0.53	L2=5.14
R3	258	W3=0.66	L3=4.99
R4	351	W4=0.75	L4=5.32
R5	299	W5=1.09	L5=4.78
R6	481	W6=1.14	L6=4.27

The phase shifter used in the model is based on broadband aperture coupling technique. The analysis of the Phase shifter is done using the conventional approach done for coupled microstrip line- treating it as a four port device with two of its ports open circuited[2].Previous research has been done on creating microwave phase shifters. However these works have often been non usable on high bandwidth equipment. Edge coupled and Schiffman phase shifters are popular but they suffer from issues such as high fabrication costs and difficulties. The design proposed in [2] has been optimal because of its relative ease to design for own purpose and high bandwidth. Three major steps have been utilized in creating this broadband phase shifter. The first step is to mathematically solve for the even and odd mode characteristic impedances for the proposed ellipses. After solving for them, an elliptical integral of the first kind was calculated for finding out the dimensions of the elliptical slots.

This calculation was complex and has been succinctly described in [2]. The length calculation however is a relatively straightforward process with quarter wave formula being used at the resonant frequency of design. The modelling of these patches is the next major step. Using the calculated dimensions, we cut two nearly circular slots with the size of length 7.9mm and minor axis of 3.2mm in the ground plane. The ground was then sandwiched between two substrates of thickness 0.508mm and dielectric constant 2.2. Above and below the slots, we designed elliptical patches of the same length (7.9 mm) and minor axis 2.41mm.

The two elliptical patches placed below the lower substrate were joined using a microstrip line of characteristic impedance 50 ohms. The topmost patches were joined to the output port and the Wilkinson divider output respectively. The last step we undertook was to optimize the dimensions of both the slots and the elliptical patches to give us a 90 degree phase shift in the lower X band between 8 to 9 GHz. Fine tuning the dimensions was a largely cumbersome process and was solved using trial and error.

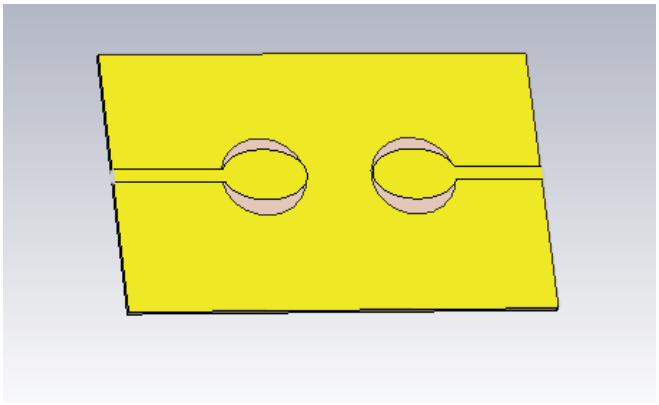


Fig. 2. Phase Shifter

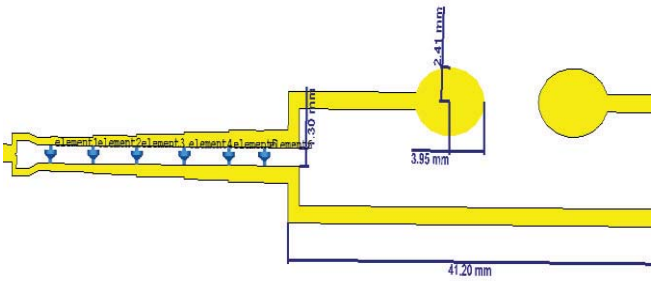


Fig. 3. Wilkinson power splitter and phase shifter combined

The output of the phase shifter and the reference line were fed to the slotted antipodal Vivaldi antenna The Vivaldi antenna was designed using a substrate of dielectric constant 2.98 and thickness 1.5 mm. Using standard design equations in [4], we were able to model the curve as an exponential taper of the type Y=Ae^(Rx). The width and length of the antennas were calculated using the lower limit of the frequency of operation and material parameters like dielectric constant of the substrate, width of the substrate, etc. Our Vivaldi antenna has a width of around 83 mm and a length spanning 91 mm. The microstrip feed line that was tapered was designed for a characteristic impedance of 50 ohms and its width can be calculated easily using software such as Txline or solving equations given in [4].

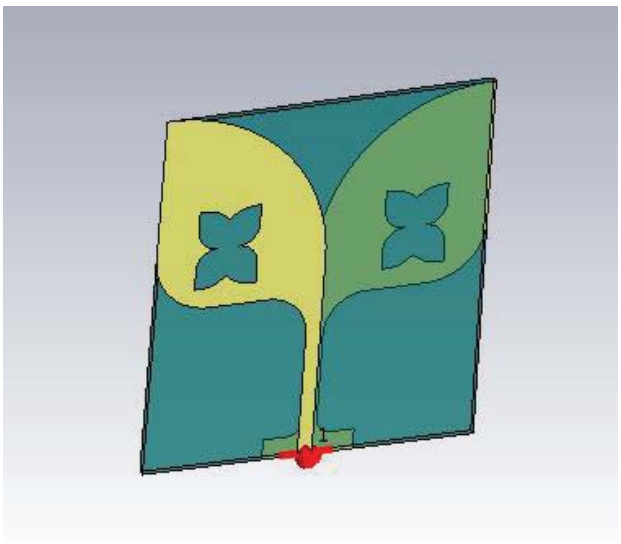


Fig. 4. Antipodal Vivaldi Antenna

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:00:44 UTC from IEEE Xplore. Restrictions apply.

The curves were optimized using a heuristic approach and the changes in the S parameters in CST Studio Suite 2019 were observed to create a broadband and compact antenna design. After the design of the radiators as well as the balun in the Vivaldi antenna, we moved on to the procedure of creating slots to enhance bandwidth. The antenna was already operating with great characteristics at 8.5 GHz but the bandwidth was only 1.2 GHz. Using techniques suggested in [4], we aimed to create floral pattern based slots in the antenna to effectively increase its bandwidth.

The slots introduced in the antenna were constructed by trial and error. The trial and error approach considered the overall impedance of the system. The slots were first designed as rectangular slots by fixing one end and using arbitrary lengths to define the rectangle and then tapering the ends to create a floral structure [4]. To achieve circular polarization, both the antennas must be kept in such a way that their phase centers are close to each other, but factors like size and the space required to prevent coupling counters it. So, for optimization, the antennas are kept at right angles to each other.

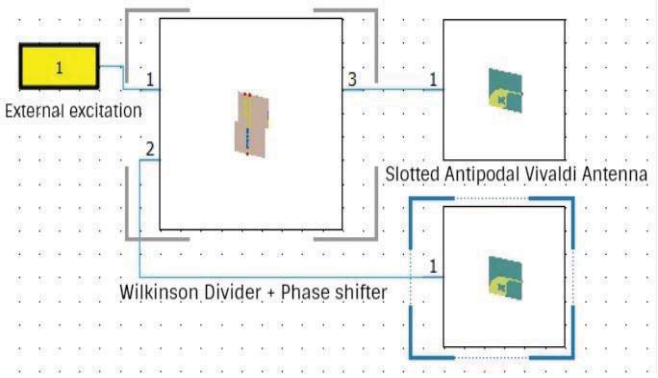


Fig. 5. Schematic of the system in CST design studio

III. RESULTS

At first, we analyzed the results of Wilkinson power divider, the phase shifter and the antipodal Vivaldi antenna individually followed by the analysis of the system. The power divider showed maximally flat response for S21 and S31 while giving great isolation (~30 dB) between the two ports.

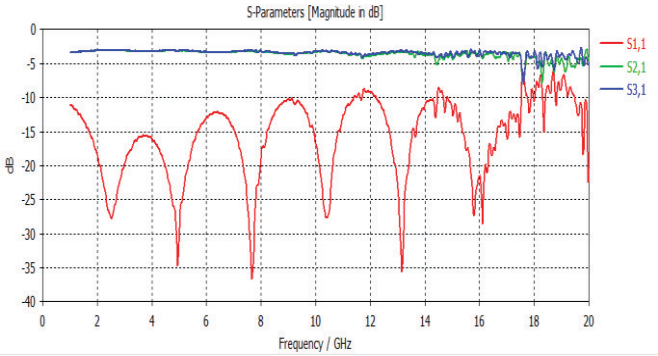


Fig. 6. S11, S21 and S31 of Wilkinson Power splitter

The Phase shifter in the 8-9 GHz band gave 90 degrees phase difference with the reference line. The slotted antipodal Vivaldi antenna showed high bandwidth.

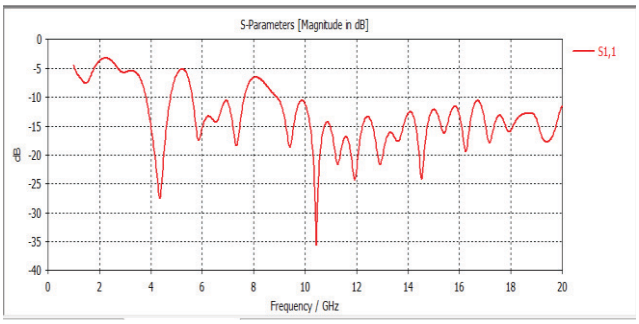


Fig. 7. Return loss of Antipodal Vivaldi Antenna

The combined array's frequency response showed the return loss below 10 db and the axial ratio below 1db for the frequency range of 8-9 GHz. Performance peaked for 8.57 GHz which showed -25 db return loss and an axial ratio of 0.098db.

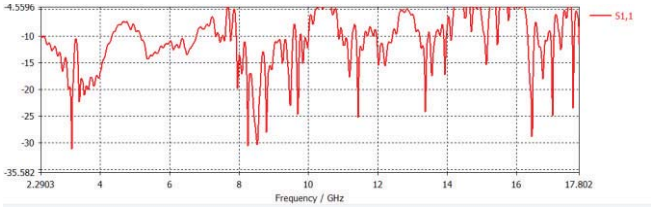


Fig. 8. Return loss of the system

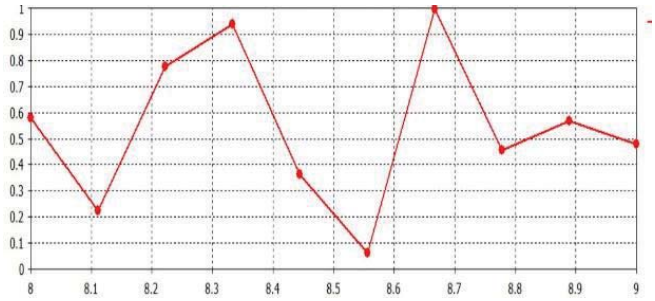


Fig. 9. Axial Ratio v/s Frequency

IV. CONCLUSION

The design of a circularly polarized antipodal Vivaldi antenna array has been presented. This array with CP gives a higher probability of successful link at even higher frequencies than presented before and eliminates the line of sight problem of linear polarization at frequencies above 2.5GHz. Vivaldi antenna in itself has been highly miniaturized and is extremely portable while giving a bandwidth of more than 7 GHz. A high isolation six stage Wilkinson divider has been employed to give maximal isolation of more than 30 dB while not being very loss. An ultra-wideband phase shifter has been designed and optimized to work at low X band frequencies to generate a phase shift of 90 degrees in the feed of the Vivaldi array and the system gives axial ratios (0.098 dB) that are much lower than those in previous literature especially in the target lower X band range.

REFERENCES

[1] Narbudowicz, A., John, M., Bao, X., & Ammann, M. J. (2012). Vivaldi array for generation of UWB circular polarization. Proceedings of the 2012 IEEE International Symposium on Antennas and Propagation. doi:10.1109/aps.2012.6348767

[2] Abbosh, A. M. (2007). Ultra-Wideband Phase Shifters. IEEE Transactions on Microwave Theory and Techniques, 55(9), 1935–1941. doi:10.1109/tmtt.2007.904051

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:00:44 UTC from IEEE Xplore. Restrictions apply.

[3] O. Kizilbey, S. Bozdemir and B. S. Yarman, "2–10 GHz multisection 2-way Wilkinson power divider with enhanced port match and isolation," 2017 IEEE 18th Wireless and Microwave Technology Conference (WAMICON), Cocoa Beach, FL, USA, 2017, pp. 1-4, doi: 10.1109/WAMICON.2017.7930250.

[4] Marek Dvorsky, Harihara S. Ganesh, S. Sadhish Prabhu, "Design and Validation of an Antipodal Vivaldi Antenna with Additional Slots", International Journal of Antennas and Propagation, vol. 2019, Article ID 7472186, 10 pages, 2019. https://doi.org/10.1155/2019/7472186

[5] D. Pozar, Microwave Engineering, 3rd ed. New York: Wiley, 2005.

[6] S. B. Cohn, "A Class of Broadband Three-Port TEM-Mode Hybrids," in IEEE Transactions on Microwave Theory and Techniques, vol. 16, no. 2, pp. 110-116, February 1968, doi: 10.1109/TMTT.1968.1126617.

[7] Guangyou, Fang & Sato, Motoyuki. (2002). Optimization of Vivaldi antenna for deming by GPR.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:00:44 UTC from IEEE Xplore. Restrictions apply.

2021 2nd International Conference for Emerging Technology (INCET)
Belgaum, India. May 21-23, 2021

Detection of Pulsars by Classical Machine Learning Algorithms

Deepanshu Beniwal
Department of Information Technology
Delhi Technological University
Delhi, India
beniwaldeepanshu442@gmail.com

Abhishek Roy
Department of Information Technology
Delhi Technological University
Delhi, India
royabhishek250@gmail.com

Himanshu Yadav
Department of Information Technology
Delhi Technological University
Delhi, India
h4himanshuyadav@gmail.com

Anamika Chauhan
Department of Information Technology
Delhi Technological University
Delhi, India
anamika@dtu.ac.in

2021 2nd International Conference for Emerging Technology (INCET) | 978-1-7281-7029-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/INCET51464.2021.9456250

***Abstract*—This paper compares the classical machine learning algorithms to classify a pulsar (rapidly spinning neutron star) in the HTRU-2 dataset [6]. This dataset was gathered during the survey, called the High Time Resolution Universe (South), and is an imbalanced dataset consisting of more than 17 thousand records. Being highly imbalanced, several methodologies have been implemented (such as sampling, resampling, and feature selection) to overcome the imbalance between target classes (binary in this case) to achieve higher accuracies. The results of the proposed work show that by applying different methodologies, it is possible to overcome the class imbalance of the dataset, which in turn helps the classical machine learning algorithms to obtain promising accuracies.**

Keywords—neutron star, pulsar, general relativity, gravitational waves, radio telescopes, binary classification, supervised machine learning, astronomy, sampling, feature selection

I. INTRODUCTION

Pulsars are quickly spinning neutron stars in a nutshell, having a diameter of about 19-20 km across. If the mass of our Sun is considered to be 1 Solar Masses, then the masses of Pulsar range between 1.18 and 1.97 Solar Masses, but most Pulsars have a mass of about 1.35 Solar Masses. In 1967, Jocelyn Bell was a postgraduate student when she was building a radio telescope. Bell noticed a noise pattern in the data recorded by their telescope, which they find unfamiliar. After many nights of analyzing those noise patterns, they finally get to know that the noise was coming from an astronomical body. She had discovered the first known pulsar. The rapid rotation of pulsars paired with their intense magnetic fields generates twin beams of energy away from their surface. As the star rotates, these beams get detected as a flash of increased brightness or pulse from Earth. This pulse carries the light of different spectrums such as visible, radio, and x-rays.

Pulsars are the leftover cores of massive stars that exploded in a supernova. Some pulsars spin with periods of a few milliseconds (called millisecond pulsar) to a couple of minutes. Some of them form a binary system with a nearby star. In rare cases, two pulsars or neutron stars get close enough such that they start orbiting each other and are known to be of very scientific importance. While orbiting each other, the neutron stars generate continuous ripples in the fabric of space-time, thus causing what we call gravitational waves, which according to cosmologists, has confirmed an important prediction of the general theory of relativity proposed by Albert Einstein in 1916.

Modern telescopes used to detect these bizarre objects produce terabytes of data per observation, pushing supercomputers to their limits and emerging as a problem for astronomers. Analyzing such data requires astronomers to be able to think computationally to solve problems. Thus, automation in detection and classification through machine learning is needed with a requirement of optimization to permit efficient and real-time throughput processing of data. The proposed work investigates the utility of supervised Machine learning in pulsar detection from the data fed into them.

The proposed paper is organized in the following manner: In Section-II, a summary of the studies previously done under pulsar detection, along with what new we're proposing in our work, is presented. Section-III explains the working methodology. In Section-IV, Experiments with setup and dataset variables description are presented. Results of the experiments are presented in Section-V. Based on the results obtained in the previous section, conclusions are made in Section-VI.

II. LITERATURE REVIEW

Researches in the field of astronomy have seen significant growth over the past couple of years. One of the reasons for this is the large amount of data generated by modern-day telescopes for every single observation they make, which astronomers have found quite overwhelming, as it is nearly impossible to deal with such an amount of data manually. Thus, computer scientists and algorithm experts who already have the required experience of handling such data have shown interest in automating astronomy-related problems. Under astronomy, the detection of pulsars is one such big problem. Pulsars are a rare type of stars and since they're rare, detecting them through petabytes of data generated is not an easy task.

Researchers and algorithm experts around the world have contributed some of their work towards this area of astronomy. Several Machine Learning and Artificial

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:48:05 UTC from IEEE Xplore. Restrictions apply.

Intelligence models are putting forward to help astronomers automating this task.

A Novel AI system called PICS or the Pulsar Image-based Classification System is introduced by [1], where pulsar detection is made possible through image pattern recognition with the help of a 9000 neurons-based deep-learning model. A survey pipeline called the PALFA or the Pulsar Arecibo L-Band Feed Array has integrated this system and, to date, has successfully discovered six new pulsars.

In [2], the authors have worked on pulsar classification with very few features, primarily focusing on feature selection procedures. They proposed two feature selection algorithms called GS (Grid Search) and RFE (Recursive Feature Elimination) and found promising recall and false-positive rates.

[3] Developed a novel approach for pulsar classification by extracting deep-features in pulsar candidate's image domain with DCGAN (deep convolutional generative adversarial networks) and by training a classifier developed using Multi-Layer Perceptron with PILAE (pseudo-inverse learning auto-encoder) algorithm.

Sometimes either due to interstellar medium or due to a farther remote location from the Earth, the signals received from a pulsar aren't strong enough to be used by traditional pulsar detection models based on FFT (Fast Fourier Transform) spectrum search. Thus, a different approach is required to overcome this weak strength of the signal. One such work is introduced by [4], where the authors have combined Bi-Spectrum (higher-order spectrum) and a deep convolutional neural network to detect weak pulsar signals. Their method involved a non-uniform sampling strategy and an extra de-noising process giving their model an accuracy greater than 90%.

The work done by [5] involves a machine learning pipeline to detect pulsars on the HTRU-2 dataset. They proposed two different strategies to overcome unbalanced data. By implementing random resampling methods with K-Fold Cross Validation (as the sampling method) and Correlation (as the feature selection method), their model gave a good insight into how resampling techniques can help overcome the unbalanced data. Their findings have inspired our work. Thus, our work can be considered an extension or a branching to their work. In our proposed work, we've implemented various resampling techniques and have introduced Stratified K-Fold Cross-Validation along with the ordinary K-Fold Cross-Validation as a sampling method. We've introduced ANOVA F-Test (as the feature selection method) and have used Accuracy and F1-Score metrics as a measure to find results through Experiments.

III. Methodology

Figure-1 represents the pipeline of the proposed work. Each component of the methodology is discussed below.

A. Data Analysis

Data Analysis is the first step in the proposed pipeline. Under this step, the attributes and properties such as the number of rows & columns, different data types (like int and float), and null values are analyzed in the dataset. Statistical measures such as mean and standard deviation of the dataset are also analyzed.

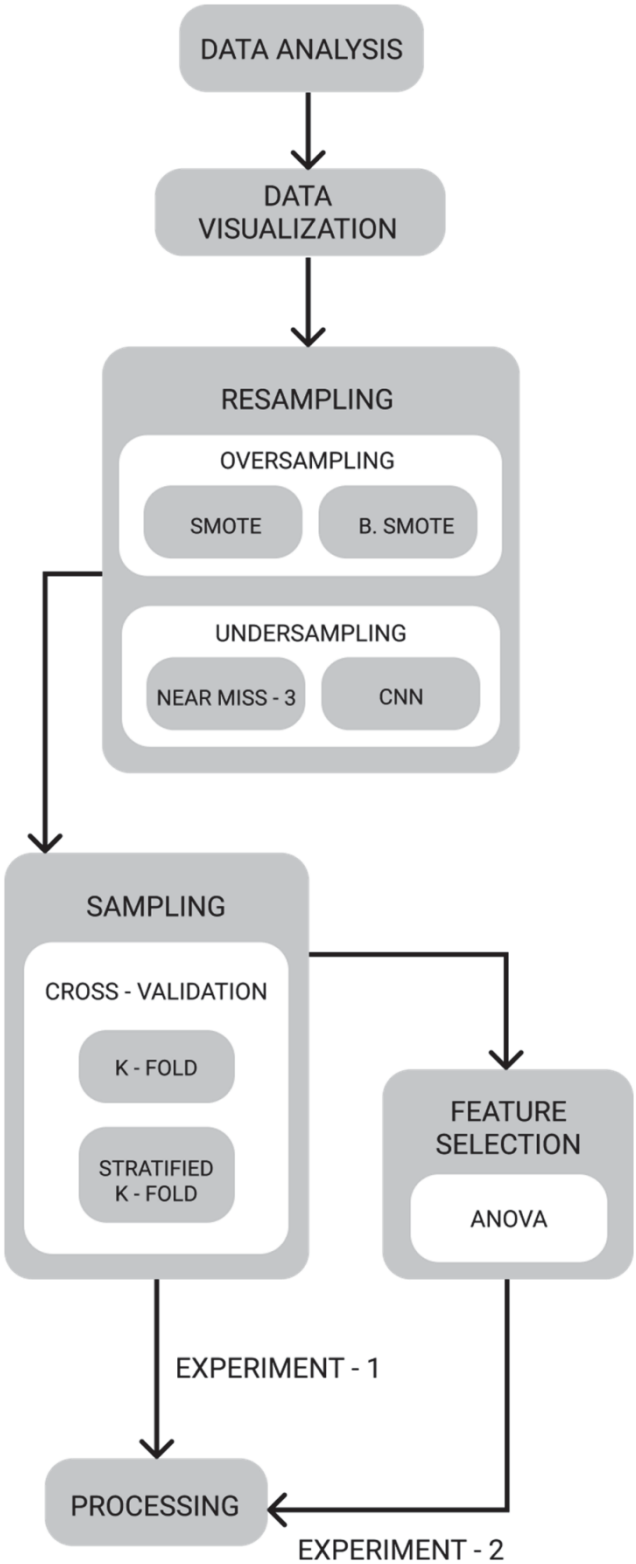


Fig. 1. Pipeline of the proposed work

B. Data Visualization

In machine learning, data visualization is used to understand data or information in a visual format such as a bar chart or a graph. It is an essential step in the exploration of a dataset in terms of class distribution. Data Visualization makes it easier to identify any underlining patterns, outliers, and class imbalances within a dataset.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:48:05 UTC from IEEE Xplore. Restrictions apply.

C. Resampling

Resampling (also known as Monte Carlo Estimation) refers to the techniques, which efficiently use data samples in order to improve the accuracy metric and assists in quantifying the uncertainty of the estimate. The two most common types of resampling techniques, oversampling and under-sampling, are discussed below.

1) Over-sampling: Oversampling is a type of resampling technique that deals with datasets having an imbalance between the target classes by generating synthetic instances for minority classes until a more balanced distribution is attained. The proposed work involves two unique oversampling methods, Smote [8] and Borderline Smote [9].

2) Under-sampling: Under-sampling is another type of resampling technique (opposite of oversampling) that deals with datasets where there is an imbalance between the target classes by discarding pre-existing instances of the majority class until a more balanced distribution is attained. This work involves two under-sampling methods, namely, Near Miss-3 and Condensed Nearest Neighbors [11].

D. Sampling

Sampling is a statistical approach concerned with choosing the individual observation, making it easier to draw statistical inferences about the data. A type of sampling method, called Cross-Validation, used for the validation of models, is described below:

1) Cross-Validation: It is a sampling technique used to train models by the subset of the dataset, which in turn gets evaluated by the complementary subset of the same dataset. In our work, we've implemented two Cross-Validation techniques, namely K-Fold Cross-Validation and Stratified K-Fold Cross-Validation.

E. Feature Selection

It is a means of selecting the most important features of a dataset based on different methods such as Pearson correlation, recursive feature elimination, and more. Restricting the number of features helps in reducing irrelevant or redundant features to be used in the training of the model, which in turn often improves the model's accuracy.

The proposed work uses the ANOVA F-test (analysis of variance f-test) [10] as the basic criteria for feature selection, which is a parametric statistical hypothesis test for determining whether the means from two or more data samples come from the same distribution or not.

F. Processing

Processing is the last step in the pipeline yet the most important. It forms the basis for comparison between results obtained by different classical ML algorithms [12] [13]. The classical algorithms used in the proposed work are listed as follows:

- Logistic Regression
- K – Nearest Neighbors (KNN)
- Decision Trees
- Random Forest
- Naive Bayes

IV. Experiments

A. Dataset

This work uses the HTRU-2 data-set consisting of samples of pulsar candidates gathered through the High Time Resolution Universe Survey and are all verified by human analysts.

This dataset consists of 17,898 instances, of which 16,259 are false instances originated by RFI/noise belonging to the majority negative class with a class label of 0 (negative), and 1,639 are real pulsar instances belonging to the minority positive class with a class label of 1 (positive).

For every pulsar candidate in the dataset, there are a total of eight statistical variables, half of which belong to the DM-SNR curve profile, and the other half belong to the integrated pulse profile. The statistical variables used are mean, excess kurtosis, standard deviation, and skewness for both the profiles.

This section explains the followings terms:

- Integrated pulse profile
- DM-SNR curve
- Skewness
- Kurtosis

1) Integrated Pulse Profile: The individual pulse or blip of increased brightness from a pulsar lacks a constant shape or amplitude in nature, making them not very useful for identifying pulsars uniquely. Since the independent emissions from a pulsar are different in nature, astronomers stack up multiple recorded pulses thus, creating an average integrated pulse profile to characterize a particular pulsar, as shown in figure 2 [7].

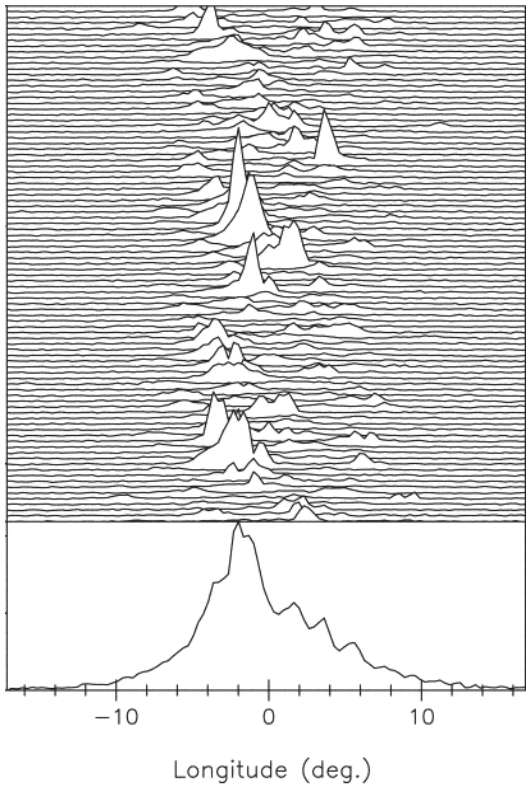


Fig. 2. Average integrated pulse profile

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:48:05 UTC from IEEE Xplore. Restrictions apply.

2) DM-SNR Curve: Another uncertainty that gives rise to a new profile is that the pulses arrive at different times across different radio frequencies. Ionized interstellar medium gives rise to this delay from frequency to frequency (figure 3 [7]) and, this phenomenon is called dispersion. To overcome this effect, Astronomers use several approaches to compensate for the shape of the delay. Even after that, there is always some uncertainty associated with the fit, expressed by the DM-SNR curve, as depicted in figure 4 [7].

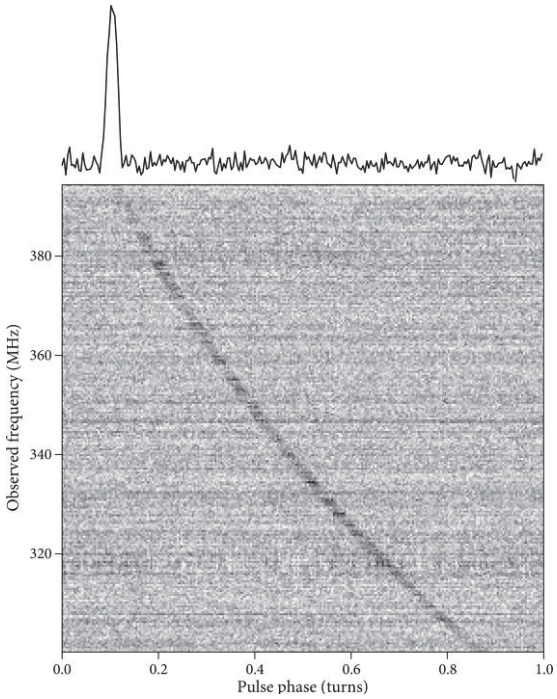


Fig. 3. Delay from frequency to frequency

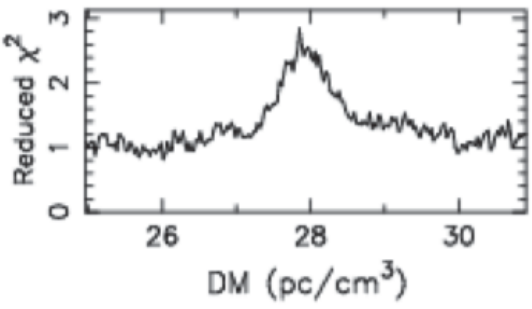


Fig. 4. The DM-SNR curve

3) Skewness: The Skewness of a probability distribution measures the asymmetry of the probability distribution about its mean. Skewness can have values zero, positive, negative, or undefined. The below figure shows the general graphical representation of Skewness:

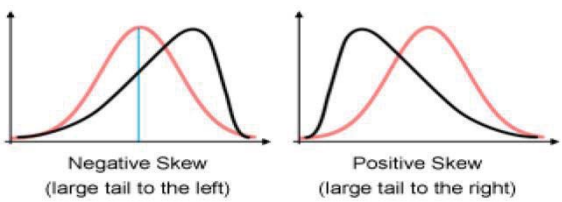


Fig. 5. General graphical representation of Skewness

4) Kurtosis: Kurtosis measures the "peakedness" of the distribution and outlines the shape of the same. The below figure shows the general graphical representation of Kurtosis:

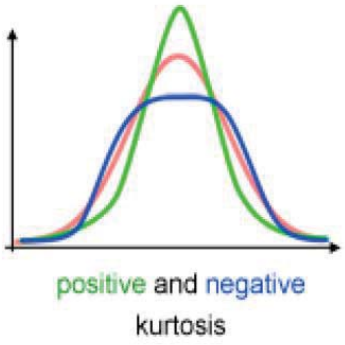


Fig. 6. General graphical representation of Kurtosis

B. Experimental Setup

Both experiments were carried out using five variants of the HTRU-2 dataset in the view of two under-sampling and two over-sampling algorithms in order to repress the imbalance of target classes in the dataset.

1) Dataset naming Convention:

a) ORIGINAL: The original or genuine HTRU-2 dataset without any modification.

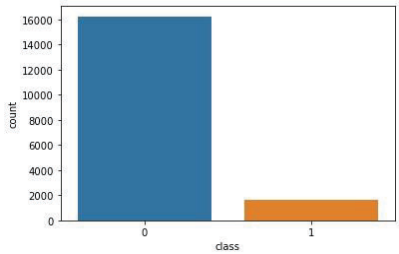
b) SMOTE: Dataset generated using the Smote oversampling algorithm over the genuine dataset.

c) B. SMOTE: Dataset generated using the Borderline Smote oversampling algorithm over the genuine dataset.

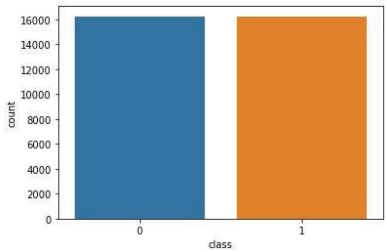
d) NM-3: Dataset generated using the Near Miss-3 undersampling algorithm over the genuine dataset.

e) CNN: Dataset generated using the Condensed Nearest Neighbors undersampling algorithm over the genuine dataset.

2) Visual Class Distribution of all Datasets

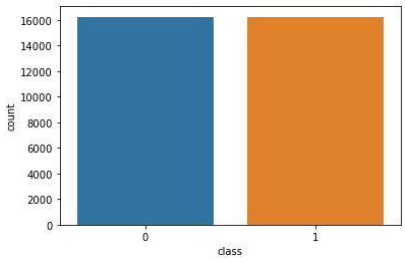


(a)

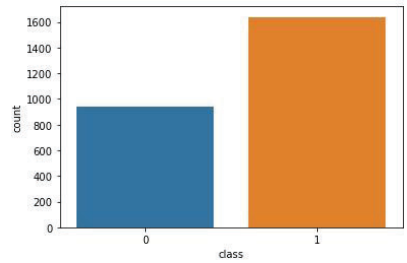


(b)

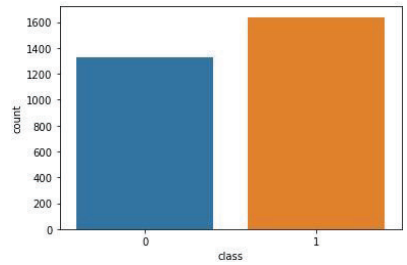
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:48:05 UTC from IEEE Xplore. Restrictions apply.



(c)



(d)



(e)

Fig. 7. Visual Class Distribution of all five datasets: (a) ORIGINAL Dataset, (b) SMOTE Dataset, (c) B. SMOTE Dataset, (d) NM-3 Dataset, and (e) CNN Dataset.

C. Experiment-1

This experiment is performed as per the five datasets mentioned above without feature selection with K-Fold Cross Validation for sampling. The dimensions of datasets used in this experiment are shown in Table-I below:

TABLE I. PROPERTIES OF DATASET USED IN EXPERIMENT-1

Dataset Name	Class Distribution	Dimensions
ORIGINAL	[16259, 1639]	[17898, 9]
SMOTE	[16259, 16259]	[32518, 9]
B. SMOTE	[16259, 16259]	[32518, 9]
NM-3	[1639, 941]	[2580, 9]
CNN	[1639, 1330]	[2969, 9]

D. Experiment-2

This experiment is performed as per the five datasets mentioned above with only six important features, obtained using the ANOVA F-Test Algorithm with Stratified K-Fold Cross Validation for sampling.

The six important features obtained by ANOVA F-Test are:

1) *Mean of the IPP*
2) *Excess kurtosis of the IPP*
3) *Excess kurtosis of the DCP*
4) *The skewness of the IPP*
5) *The standard deviation of the DCP*
6) *Mean of the DCP*

Where,

- IPP stands for *Integrated Pulse Profile*
- DCP stands for *DM-SNR Curve Profile*

The dimensions of datasets used in this experiment are shown in Table II below:

TABLE II. PROPERTIES OF DATASET USED IN EXPERIMENT-2

Dataset Name	Class Distribution	Dimensions
ORIGINAL	[16259, 1639]	[17898, 7]
SMOTE	[16259, 16259]	[32518, 7]
B. SMOTE	[16259, 16259]	[32518, 7]
NM-3	[1639, 941]	[2580, 7]
CNN	[1639, 1330]	[2969, 7]

V. RESULTS

A. Experiment-1

The experiment was performed using the K-Fold Cross Validation sampling method with k (number of folds) = 10. Table-3 below shows the results when accuracy is taken into account as a measure.

TABLE III. EXPERIMENT-1 RESULTS USING ACCURACY AS A MEASURE

Classifiers	Datasets				
	ORIGINAL	***SMOTE***	***B. SMOTE***	***NM-3***	***CNN***
Naive Bayes	0.94	0.90	0.86	0.78	0.84
Logistic Regression	0.97	0.93	0.93	0.82	0.88
KNN	0.97	0.95	0.96	0.61	0.80
Random Forest	0.97	0.97	0.98	0.78	0.87
Decision Trees	0.96	0.95	0.97	0.75	0.81

The above results clearly show that neither oversampling nor under-sampling performed as expected. The original dataset with a high imbalance between target classes performs better than any other dataset. Apart from this, under-sampling has performed poorly in comparison to oversampling. Below is a trend of datasets from best to worst (based on Accuracy):

ORIGINAL > B. SMOTE > SMOTE > CNN > NM-3

TABLE IV. EXPERIMENT-1 RESULTS USING F1-SCORE AS A MEASURE

Classifiers	Datasets				
	ORIGINAL	***SMOTE***	***B. SMOTE***	***NM-3***	***CNN***
Naive Bayes	0.67	0.80	0.75	0.59	0.52
Logistic Regression	0.84	0.87	0.82	0.63	0.54
KNN	0.79	0.84	0.85	0.57	0.52
Random Forest	0.83	0.93	0.93	0.62	0.54
Decision Trees	0.75	0.86	0.89	0.62	0.54

Table-IV shows the results when f1-score acts as a measure.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:48:05 UTC from IEEE Xplore. Restrictions apply.

The results in Table-IV show that oversampling is a better approach for lesser misclassification (i.e., lesser false positives and false negatives) than others as oversampling results in the highest F1-Scores (since it is the weighted harmonic mean of precision and recall). Under-sampling, on the other hand, has a poor F-1 score which makes it clear why even in the results of Table-III, under-sampling had poor accuracies. The original dataset stands between over-sampling and under-sampling. Below is a trend of datasets from best to worst (based on F1-Score):

SMOTE > B. SMOTE > ORIGINAL > CNN > NM-3

Common findings for Experiment-1:

- KNN performs worst, whereas Logistic Regression performs best in under-sampling.
- Naive Bayes performs worst, whereas Random Forest performs best in oversampling.

B. Experiment-2

The experiment was performed using the Stratified K-Fold Cross Validation sampling method with k (number of folds) = 10. Table-5 below shows the results when accuracy is taken into account as a measure.

TABLE V. EXPERIMENT-2 RESULTS USING ACCURACY AS A MEASURE

Classifiers	Datasets				
	ORIGINAL	*SMOTE*	*B. SMOTE*	*NM-3*	*CNN*
Naive Bayes	0.94	0.90	0.84	**0.80**	0.84
Logistic Regression	0.97	0.94	0.93	**0.86**	**0.89**
KNN	0.97	0.95	0.93	**0.76**	**0.83**
Random Forest	0.97	0.97	0.96	**0.84**	**0.88**
Decision Trees	0.96	0.95	0.95	**0.79**	**0.83**

From the above results, one can see that using feature selection didn't improve the classification by any means. The reason why under-sampling performed better (marked in bold) in this Experiment, compared to that in Experiment-1 is, because of the use of Stratified K-Fold Cross Validation, instead of normal K-Fold Cross Validation, which found use in Experiment-1.

Comparing to the results of Experiment-1, the accuracies of only under-sampling are improved. The rest is almost the same as Experiment-1.

Below is a trend of datasets from best to worst (on the basis of Accuracy):

ORIGINAL > SMOTE > B. SMOTE > CNN > NM-3

Table-VI shows the results when f1-score acts as a measure.

TABLE VI. EXPERIMENT-2 RESULTS USING F1-SCORE AS A MEASURE

Classifiers	Datasets				
	ORIGINAL	*SMOTE*	*B. SMOTE*	*NM-3*	*CNN*
Naive Bayes	**0.74**	**0.89**	**0.83**	**0.82**	**0.84**
Logistic Regression	**0.87**	**0.94**	**0.93**	**0.89**	**0.90**
KNN	**0.84**	**0.95**	**0.92**	**0.80**	**0.84**
Random Forest	**0.88**	**0.97**	**0.95**	**0.87**	**0.89**
Decision Trees	**0.81**	**0.95**	**0.94**	**0.84**	**0.85**

We already know that feature selection didn't improve our model in any way. The F1-Scores of Table-6 above implies that Stratified K-Fold Cross Validation alone has reduced the misclassifications previously made in Experiment-1 to a significant amount as the F1-Scores in Table-VI are way better (marked in bold) than that they were in Experiment-1.

Oversampling has performed the best here too, but the interesting part is that under-sampling has performed better than the original dataset, which was opposite in the case of F1-Scores of Experiment-1, where under-sampling performed poorly compared to the original dataset. Below is a trend of datasets from best to worst (on the basis of the F-1 Score):

SMOTE > B. SMOTE > CNN > NM-3 > ORIGINAL

Common findings for Experiment-2:

- KNN performs worst, whereas Logistic Regression performs best in under-sampling.
- Naive Bayes performs worst, whereas Random Forest performs best in oversampling.

VI. CONCLUSION

- Sampling methods such as Stratified K-Fold Cross-Validation can drastically improve the misclassifications made by the model resulting in higher F1- Scores.
- Random Forest & Logistic Regression classifiers can always be used in pulsar classification as they perform the best in over-sampling and under-sampling, respectively.
- Feature Selection doesn't seem to improve the model accuracy as all the features hold some information crucial to classify candidates as pulsars.
- As being highly imbalanced, there is always a need for sampling and resampling methods in order to repress the imbalance of target classes in the HTRU-2 dataset.

REFERENCES

[1] W.W. Zhu, A. Berndsen, E.C. Madsen, M. Tan, I.H. Stairs, A. Brazier, P. Lazarus, R. Lynch, P. Scholz, K. Stovall, S.M. Ransom, S. Banaszak, C.M. Biwer, S. Cohen, L.P. Dartez, J. Flanigan, G. Lunsford, J.G. Martinez, A. Mata, M. Rohr, A. Walker, B. Allen, N.D.R. Bhat, S. Bogdanov, F. Camilo, S. Chatterjee, J.M. Cordes, F. Crawford, J.S. Deneva, G. Desvignes, R.D. Ferdman, P.C.C. Freire, J.W.T. Hessels, F.A. Jenet, D.L. Kaplan, V.M. Kaspi, B. Knispel, K.J. Lee, J. van Leeuwen, A.G. Lyne, M.A. McLaughlin, X.Siemens, L.G. Spitler, and A. Venkataraman: SEARCHING FOR PULSARS USING IMAGE PATTERN RECOGNITION. The Astrophysical Journal, Volume 781, Number 2, https://iopscience.iop.org/article/10.1088/0004-637X/781/2/117/pdf

[2] Haitao Lin, Xiangru Li, Ziying Luo: Pulsars Detection by Machine Learning with Very Few Features. https://arxiv.org/abs/2002.08519v1

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:48:05 UTC from IEEE Xplore. Restrictions apply.

[3] Mohammed A.B. Mahmoud, Ping Guo: Learning from imbalanced pulsar data by combine DCGAN and PILAE algorithm, https://doi.org/10.1016/j.newast.2020.101561

[4] Wang, Longqi & Jin, Jing & Jiang, Yu & Shen, Yi. (2019). A Method for Weak Pulsar Signal Detection Combining the Bispectrum and a Deep Convolutional Neural Network. The Astrophysical Journal. 873. 17. 10.3847/1538-4357/ab0308.

[5] Alexander Ylnner Choquenaira Florez, Braulio Valentin Sanchez Vinces, Diana Carolina Roca Arroyo, and Josimar Edinson Chire Saire Patrıcia Batista Franco {alexanderchf, braulio.sanchez, dianaroca, jecs89, patricia.franco}@usp.br: Machine Learning Pipeline for Pulsar Star Dataset, https://arxiv.org/pdf/2005.01208.pdf

[6] R. J. Lyon, B. W. Stappers, S. Cooper, J. M. Brooke, J. D. Knowles, Fifty Years of Pulsar Candidate Selection: From simple filters to a new principled real-time classification approach, Monthly Notices of the Royal Astronomical Society 459 (1), 1104-1123, DOI: 10.1093/mnras/stw656

[7] James J. Condon and Scott M. Ransom: Essential Radio Astronomy, https://www.cv.nrao.edu/~sransom/web/Ch6.html#S1

[8] N. V. Chawla, K. W. Bowyer, L. O. Hall, W. P. Kegelmeyer, SMOTE: Synthetic Minority Over-sampling Technique, Journal Of Artificial Intelligence Research, Volume 16, pages 321-357, 2002, DOI: 10.1613/jair.953, https://arxiv.org/abs/1106.1813

[9] Han H., Wang WY., Mao BH. (2005) Borderline-SMOTE: A New Over-Sampling Method in Imbalanced Data Sets Learning. In: Huang DS., Zhang XP., Huang GB. (eds) Advances in Intelligent Computing. ICIC 2005. Lecture Notes in Computer Science, vol 3644. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11538059_91

[10] Max Kuhn, Kjell Johnson: FEATURE ENGINEERING AND SELECTION: A Practical Approach for Predictive Models, 2019, Page 242

[11] P. Hart, "The condensed nearest neighbor rule (Corresp.)," in IEEE Transactions on Information Theory, vol. 14, no. 3, pp. 515-516, May 1968, doi: 10.1109/TIT.1968.1054155.

[12] M. Wadera, M. Mathur and D. K. Vishwakarma, "Sentiment Analysis of Tweets- A Comparison of Classifiers on Live Stream of Twitter," 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2020, pp. 968-972, doi: 10.1109/ICICCS48265.2020.9121166.

[13] P. Meel, M. Singhal, M. Tanwar and N. Saini, "Predicting Flight Delays with Error Calculation using Machine Learned Classifiers," 2020 7th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 2020, pp. 71-76, doi: 10.1109/SPIN48934.2020.9071159.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:48:05 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021)
IEEE Xplore Part Number: CFP21K74-ART; ISBN: 978-0-7381-1327-2

Detection Of Skin Cancer Using Neural Architecture Search with Model Quantization

Ritabrata Maiti
Computer Science and Engineering
Delhi Technological University
New Delhi, India
ritabratamaiti_2k17co268@dtu.ac.in

Priyanka Agarwal
Computer Science and Engineering
Delhi Technological University
New Delhi, India
priyankaagarwal_2k17co243@dtu.ac.in

Ragyayee Ravinder Kumar
Computer Science and Engineering
Delhi Technological University
New Delhi, India
ragyayeeravinderkumar_2k17co251@dtu.ac.in

Aruna Bhat
Computer Science and Engineering
Delhi Technological University
New Delhi, India
aruna.bhat@dtu.ac.in

2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) | 978-1-6654-1272-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICCS51141.2021.9432190

***Abstract*—Skin cancer is the fifth most common cancer across the globe. One in every three cancers detected is skin cancer. Every year, between 2 and 3 million non-melanoma skin cancers and 132,000 melanoma skin cancers occur globally each year. The worldwide number of cases of carcinoma is increasing by epidemic proportions. Catching cancer early often allows for more treatment options. It is crucial to know which type of cancer one has because it affects the treatment options and prognosis. Hence, our work is primarily motivated by the importance of early and efficient detection of skin cancer. In this paper, a novel methodology has been proposed by incorporating Neural Architecture Search and Model Quantization technique, where deep convolutional neural network is utilized to build a skin cancer classifier, which can be deployed on low compute devices. To the best of our knowledge this is the first instance of utilizing Neural Architecture Search in combination with Model Quantization for the purpose of Skin Cancer detection. The yielded results show that, from the model proposed through this methodology, the model size can be reduced without causing a significant change in the accuracy. Thus, making this type of a skin classifier a more effective tool for skin cancer detection allows it for the usage on devices with limited memory and computational capacity.**

Keywords*—Skin Cancer, neural architecture search, autoML, EfficientNets, quantization.

I. INTRODUCTION

According to the American Cancer Society,[1] About 106,110 new melanomas will be diagnosed and About 7,180 people are expected to die of melanoma in the USA. Skin cancer is the development of abnormal cells in the outermost layer of the skin. Such changes are triggered by unpaired DNA damage leading to rapid replication and formation of malignant tumors. Basal cell carcinoma (BCC), squamous cell carcinoma (SCC), melanoma, and Merkel cell carcinoma (MCC) are the main types of skin cancer. Catching cancer early often allows for more treatment options. It is crucial to know which type of cancer one has because it affects the treatment options and prognosis. Early diagnosis dramatically increases the prognosis of patients with malignant melanoma.

The initial diagnosis step of skin cancer is through the detection of suspected spots visually, either through self-examinations or professional care. If left untreated, skin cancer can even cause death. When it comes to visual detection, color, texture, size, shape of the suspected area get examined. Biopsies and imaging tests are ways for the early detection of skin cancer. The different imaging procedures include CT-scan, XR ray, and MRI. There have been attempts at skin cancer detection using CAD, machine learning algorithms, convolutional neural networks. Most of these techniques relied on the state-of-art methods available back then. However, there has been a significant improvement over the past few years for speedy diagnosis at an early stage.

While there have been previous attempts at creating Dermatologist-level classification models for diagnosing skin cancers using deep learning techniques, most of them have been focused on using pre-existing deep learning architectures[1]. Such architectures were created with a different purpose in mind: to create general-purpose visual object recognition models and perform well in large-scale image recognition datasets like Open Images and Image Net. Therefore, this paper proposes a novel methodology that

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:33:35 UTC from IEEE Xplore. Restrictions apply.

focuses on a group of methods called neural architecture search that helps in finding an effective architecture in an automated manner, hence producing satisfactory results in less time[2].NAS proves to be effective in finding a proper network structure by taking different aspects of the problem into consideration. The design process takes into account the complexity of the problem, the size of the dataset, implementation constraints, learning time, and other parameters.

In order to effectively utilize the skin cancer classifier as a medical tool, it must be converted into a less complex model. This paper also extends this approach for easy implementation in low resource devices using the concept of Quantization Aware training by reducing the size of the model while maintaining its accuracy. The size of the model gets reduced 4 times as quantized models use lower precision weights. The proposed approach proves to be effective in choosing the right architecture with reduced model size while providing more accurate results compared to previous state-of-art techniques.

The rest of the paper is organized as follows – Section II has the Background and Related Work. The Methodology is given in Section III and Experiment and Results in Section IV. Section V has the contents of Conclusion and Future Work.

II. BACKGROUND AND RELATED WORK

Over the past decades, the medical field has experienced an immense advancement in research solutions for skin lesion detection. Most of these techniques relied on the state-of-art methods available back then. However, there has been a significant improvement over the past few years for speedy diagnosis at an early stage. These techniques vary from using Computer-Aided Diagnosis systems (CAD)[3][4]Machine learning algorithms such as decision tree algorithms [5] Bayesian classifiers [6], support vector machines [7][8] to the use of Convolutional Neural Networks[9][10]. Other research involved the usage of an ensemble of deep models in place of using a single method.[11][12],feature aggregation of different models[13]. Improvements in image classification techniques were observed not only due to advancements in architecture but also in areas like optimization, regularization, activation functions, data augmentation.

Performances of different deep learning models over the years were analyzed and compared to conclude that most of these techniques were focused on using preexisting deep learning architectures. A two-stage framework consisting of FCRN and DRN is considered to be one of the first deep learning models for detection of skin cancer providing better results than the handcrafted feature models [8]. From Inception v4 CNN model[14], ResNet-152 CNN model[15], GoogleNet Inception v3 CNN model[9], An ensemble composed of DRN's, CNNs, and Fully CNNs[10] to An ensemble composed of ResNets, Densenets, and Senets[11] -such architectures were created with a different purpose in mind: to create general-purpose visual object recognition models and perform well in large scale images recognition datasets like Open Images and ImageNet. Therefore we were inspired[2] to propose a unique solution that allows for automated engineering of neural architecture, which iteratively improves and adapts in accordance with the problem domain. Neural Architecture Search focuses on finding effective architecture in an automated manner.

III. METHODOLOGY

A. Dataset and Preprocessing

For enhancing the research conducted on automatic digital diagnosis of data on dermatoscopic images the HAM10000 dataset was released. The deficiency of variety within the accessible dataset of dermatoscopic images and this problem has been tackled by the release of the "Human Against Machine with 10000 training images" - HAM10000 dataset[16]. The HAM 10000 dataset contains 10015 dermatoscopic images which can very well aid as a training set for machine learning tenacities in the field of academia. The cases in the dataset include an illustrative assortment of the vital diagnostic classes in the domain of color showing lesions.

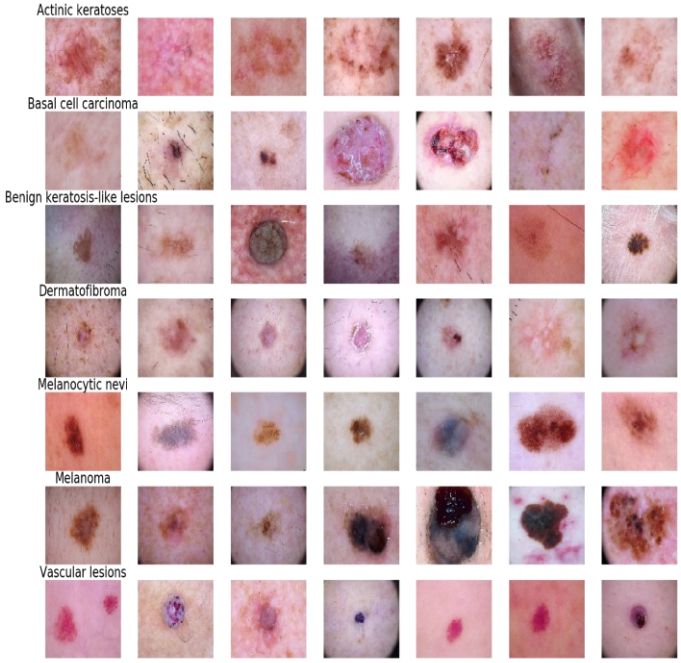


Fig.1. Skin Lesion Images from HAM10000 Dataset

The original images in the dataset were resized from 600 X 450 to 300 X 300 prior to performing the architecture search. We also normalize the dataset by dividing each pixel value with 255.0 to rescale them within the 0-1 range. We perform image data augmentation for improving the results from training. We train the model for 30 epochs with 8000 image

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:33:35 UTC from IEEE Xplore. Restrictions apply.

samples per epoch (batch size is 32). We employ random horizontal and vertical flipping when generating augmented images. Furthermore, random height and width shifts of 0.1 is also employed. A zoom range of (-0.1, 0.1) is also applied randomly for performing image augmentation. Finally, a random rotation between 0 to 360 degrees is used for the augmentation process.

B. Neural Architecture Search

In order to determine the structure of the neural network, we first have to specify the search space (E). In our Neural Architecture Search framework[17], there are primarily 2 methods for defining a search space for search deep-learning architectures with image inputs: using a pre-configured Image Classification or regression class or designing a custom block based model which defines the search space (and subsequently the architecture of the resulting neural network (a block is a step in the graph of the search-space). In our implementation, we choose the latter option. This allows us to control the size and the architecture variant in our final model.

In our approach, the search space E utilized includes every neural architecture that can be morphed from an initial architecture. The search space is tree-structured, with each child node representing a neural architecture obtained by morphing the parent node. Operations that change the number of filters in the convolutional layers increases the size of E, as compared to fixed layer width.

Given E: neural architecture search space
I: Input Data (Divided into Itrain andIval)
Cost(.): Evaluation metric function (like mean squared error)
we aim to find e*E

$$e^* = \text{Cost}(e^*, \text{Ival})$$
$$^* = L(e, \text{Itrain})$$

*is the learned parameter of e.

If n is the number of architectures and is the time taken to train each network, the time complexity of Neural Architecture Search is O(nτ). Many existing NAS techniques require a large n to reach acceptable performance, and training time for each network n is high. The NAS technique implemented in this paper reduces network morphism. Trained neural networks in E are modified into a new architecture using network morphism operations that insert a layer or add a skip-connection while keeping the architecture's original functionality the same[18]. This method finds an optimal network $e^* \in E$ which has the lowest mean squared error on the dataset. Morph operations performed here to change the number of filters in a convolutional layer make this search space larger than ones that have fixed layer widths [19]
In our customized search space, we select the EfficientNet architecture[20], followed by a max-pooling operation for two-dimensional spatial data to reduce the dimensionality to a single-dimensional vector output. We then attach a DenseBlock to the output of the max pooling layer. A DenseBlock builds a series of connected dense layers with intermediate dropout layers according to a configuration determined by input-hyperparameters. The Neural Architecture Search process controls the hyperparameters received by each block and utilizes them to configure the architecture of that block. This DenseBlock is then connected to an output node which is used for class prediction.

The input-hyperparameters received by the DenseBlock include the number of layers which will be present in this block, the number of units within each dense layer, the dropout layer for the layers and a decision variable which determines if batch normalization is to be used or not. These input parameters can be specified or tuned automatically by the Neural Architecture Search process. In our implementation, we use the bayesian optimization [autokeras] in-order to tune the parameters of the dense block.

We have used the EfficientNet architecture because of its better accuracy and efficiency, as compared to previous convolutional neural network architecture. The EfficientNet B7 architecture has been shown to achieve the state-of-the-art accuracy of 84.3% top-1 accuracy on the ImageNet dataset despite being 8.4x smaller and 6.1x faster than the best previously existing ConvNet, during inference. We choose the EfficientNet-B3 variant in our architecture in-order to reach a trade-off between model size, accuracy and training efficiency.

EfficientNets:

Convolutional Neural Networks (CNNs) are an integral part in the field of Computer Vision. They are usually developed at a fixed resource cost. One of the main issues with using CNNs is that of model scaling. Models can be scaled up or increased in size on availability of appropriate resources in order to enhance accuracy. The scaling of ResNet from ResNet-18 to ResNet-200 is made possible by consciously increasing the number of layers. The traditional way to scale a model is by increasing the width or depth of the Convolutional Neural Network, or by using a bigger input of image resolution. However, this process is time consuming and tedious and often still yields average unsatisfactory accuracy.

In 2019, Google released a paper on EfficientNets [20] which are a type of CNNs with much better accuracy. The efficiency of models is enhanced by reducing parameters and also FLOPS (Floating Point Accuracy Per Second). This is done through a compound scaling method applied on a simple baseline architecture. A grid search is performed on the model to find different relationships between the model dimensions when resource constraint is fixed. This way the scaling coefficient for dimensions is determined. The baseline architecture is developed by using AutoML to perform a Neural Architecture Search. In comparison to the traditional

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:33:35 UTC from IEEE Xplore. Restrictions apply.

CNNs, EfficientNet models usually achieve better accuracy and higher efficiency and performance.

C. Algorithm

Algorithm 1: Performing Neural Architecture Search and creating Classification Model

Input: Training Dataset D

Output: Trained neural network M for skin lesion prediction

1. Split D into train and test datasets D1 and D2 with D1 containing 9013 images and D2 1002 images.
2. Initialize Image Augmentation generators for D1, D2 as A1, A2 respectively
3. Initialize block architecture for NAS (EfficientNet-B3 architecture followed by DenseBlock)
4. Initialize hyperparameters as {training data: [1000 randomly chosen images from D1], validation data: [100 randomly chosen images from D2], maximum trials: 100, maximum epochs per trial: 5}
5. B1 = Neural_Architecture_Search(hyperparameters)
6. Recreate B1 separately and initialize with NoisyStudent weights to obtain M
7. Train M for 40 epochs using A1 generator
8. Validate and evaluate M with A2 generator
9. Save model M in appropriate format for quantization and deployment

After determining the architecture of the model by performing Neural Architecture Search, we recreate the same architecture from scratch and initialize it with the NoisyStudent weights. NoisyStudent is a semi-supervised learning approach wherein a teacher model is trained on labeled images, after which it is used to generate pseudo-labels on unlabeled images. The student network then learns to predict these pseudo-labels.

Here, we use the pre-trained weights for the EfficientNet-B3 architecture which were obtained by performing NoisyStudent training on the ImageNet dataset. We perform this step, also known as transfer-learning, in-order to strengthen the accuracy of the final model. We outline the steps we take in Algorithm 1. After we obtain the trained model M, we move on to the Quantization step.

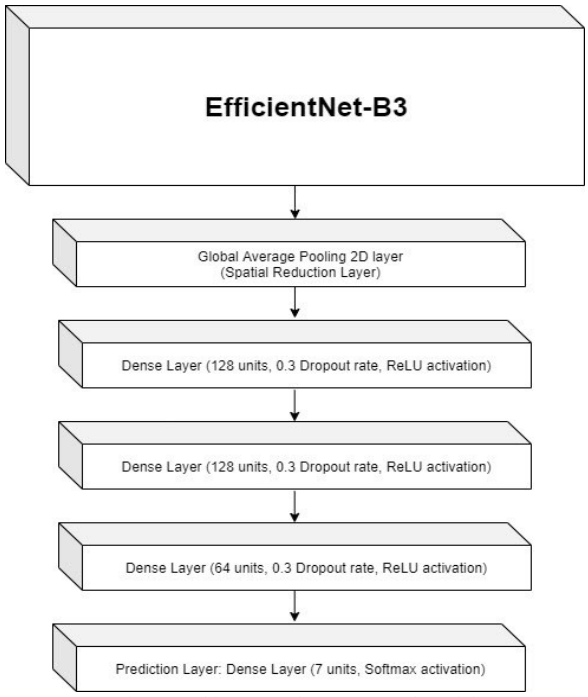


Fig.2. Architecture of our proposed model

D. Quantization

Edge devices having low computational power usually face the problem with on-device inference schemes when it comes to extensive deep learning models. With the increase in the usage of such intelligent devices, it becomes an important task to overcome this issue and propose an efficient means of optimization. To put in simple words quantization is defined as the mapping of infinite values from a large set of continuous data to a smaller set of discrete values. In this field of deep learning, a quantization scheme is to compress the neural network to a reasonable size of low bit-width numbers hence maintaining the accuracy of the model.

Two different forms of quantization can be applied to the model: Quantization Aware Training (QAT) and Post-Training Quantization. Both the techniques prove to be effective in reducing the size of the model. However, for this model, QAT is used as it provides much higher accuracy when compared to post-training quantization.[21]Quantization Aware Training is applied to the pre-trained model ensuring that the forward pass is matching both the inference time and training. QAT also plays an important role in reducing the quantization errors during training whilst maintaining the accuracy of the model. Such errors usually occur as part of the training loss. Quantize awareness can be generated either for the entire model or simply just a few parts of it.

In our architecture, rather than quantizing the entire model, we perform quantization only for specific layers. This allows us to reach a trade-off between model accuracy, size and latency. Furthermore, we focus on quantization the later layers rather than the initial layers because they enable us to achieve higher

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:33:35 UTC from IEEE Xplore. Restrictions apply.

accuracy. We quantize the final dense layers of our model before deploying the model to a TFlite backend. After quantization, we evaluate the resultant model with our test dataset (the resulting metrics can be found in section IV). After deployment to the TFLite backend we see that the accuracy persists from the previous TensorFlow format [Table III]. Furthermore, we also see that the resulting model is four times smaller than the previous model [Table IV]. Below is an example of Quantizing a simple graph consisting of one dense layer [Fig.3 and Fig.4]

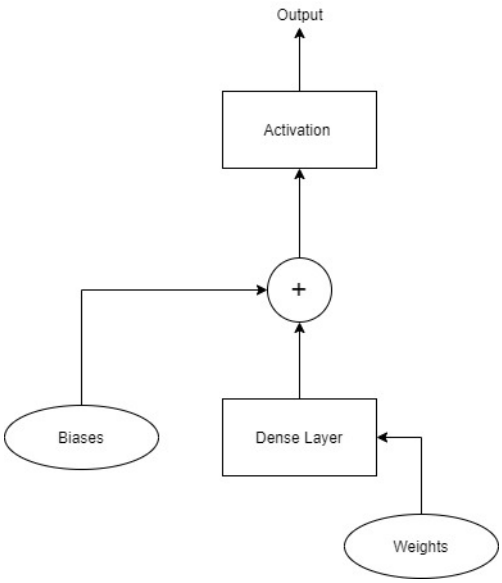


Fig.3. Simple graph consisting of one dense layer with activation (original)

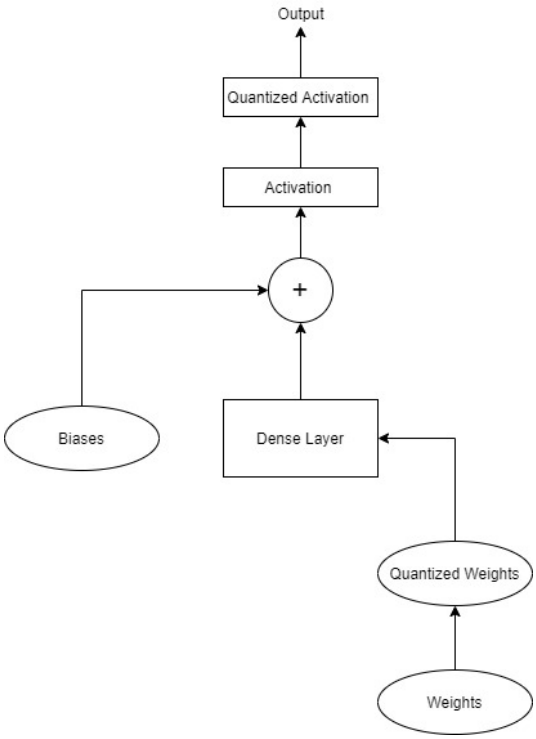


Fig.4. Quantized version of the simple graph

IV. RESULTS

We trained the model for 60 iterations and used a batch size of 32 per iteration. Our input data, configuration as described in III A, is used for training and evaluating the model. We utilise a callback function for reducing the learning rate whenever the validation accuracy plateaus. The function, upon detecting a plateau in accuracy over three iterations will reduce the learning rate by a 0.5 factor. The original learning rate is10-3 and the lower bound for the learning rate is set to 10-6.

Due to the class-imbalanced nature of the HAM10000 dataset, we use the Precision, Recall and F1-Score metrics for evaluating our proposed model.

$$Precision = \frac{\text{true positives}}{\text{true positives} + \text{false positives}} \quad (1)$$

$$Recall = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}} \quad (2)$$

$$f1\ score = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

Precision is the ratio of the true positive to the total predicted positive. Recall, on the other hand, is the ratio of the true positive to the total actual positive. F1 Score allows us to obtained a balanced score which takes into account both precision and recall. This is particularly useful when there is large class imbalance, like in the HAM10000 dataset.

Table I.
Class Wise and weighted average Precision, Recall and F1-Scores for our final model M on the Test Dataset

Lesion Type	**Metrics**		
	Precision	***Recall***	***F1 Score***
Actinic keratoses	0.88	0.74	0.81
Basal cell carcinoma	0.86	0.93	0.89
benign keratosis-like lesions	0.88	0.87	0.87
dermatofibroma	0.92	0.85	0.88
melanoma	0.78	0.71	0.74
melanocytic nevi	0.96	0.98	0.97
vascular lesions	0.95	0.95	0.95
Weighted Average	0.93	0.93	0.93

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:33:35 UTC from IEEE Xplore. Restrictions apply.

In Table I we display the precision, recall and f1-score of our trained model for each skin-lesion type, as well as the weighted average scores. This helps us in evaluating our model's performance with respect to each class of lesions, as well as the overall model performance. We observe that our model's performance varies across classifying each lesion type. This can be explained by the fact that the HAM-10000 dataset is a highly imbalanced dataset with a great degree of variation in the number of samples.

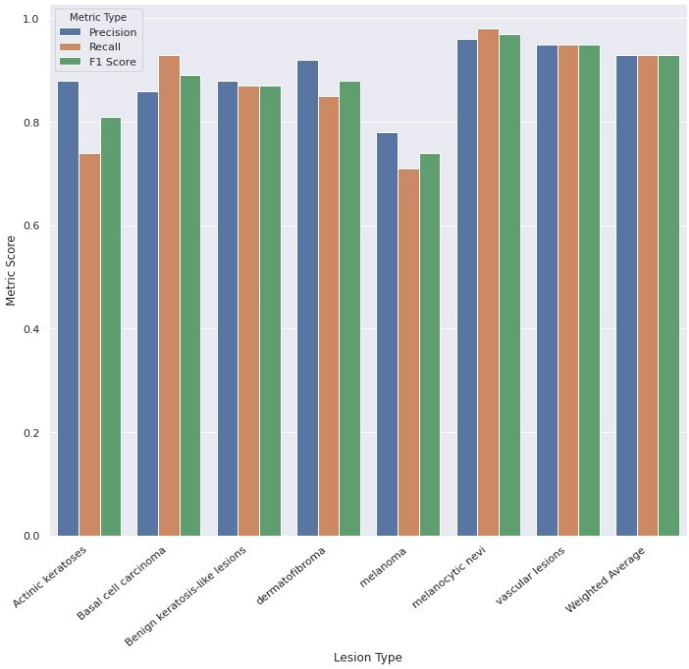


Fig.5. Class-wise metric scores for our proposed architecture

In this diagram (Fig.5), it is possible for us to visualize our model performance (along with the tabular data from Table I). This especially is useful for us to evaluate the suitability of deploying our model to end-user production applications. Since our model's precision, recall and accuracy scores are an improvement from previous techniques, we proceed to the model quantization task.

Table II.
Results of comparing our proposed architecture with some existing models

Source	Model	Avg. Recall	Avg. Precision	Avg. F1 Score
[22]	Original Densenet-121	0.49	0.36	0.30
[22]	Combining Multilevel Learnings in a DenseNet (Ensemble DenseNet Architecture)	0.88	0.76	0.82
[23]	Fine Tune MobileNet	0.89	0.83	0.83
[24]	Dilated VGG16	0.87	0.87	0.87
	Dilated VGG19	0.85	0.85	0.85
	Dilated MobileNet	0.89	0.88	0.88
	Dilated InceptionV3	0.89	0.89	0.89
Our Proposed Model	Neural Architecture Search(with EfficientNet)	0.93	0.93	0.93

In Table II we compare our proposed architecture with some of the existing models. In order to perform the comparison, we use the precision, recall and F1 metrics. We observe that our model's performance is superior to that of the preexisting architectures.

Table III.
Accuracy of the Base Model M before and after Quantization Aware Training

Accuracy of Base Model	0.93114
Accuracy of Base Model with Quantization Aware Training	0.86527

In Table III, we see that the accuracy of the base model prior-to and after its quantization is approximately 93.1% and 86.5% respectively. Thus, the cost of quantizing the model is an around 6.5% decrease in accuracy. This indicates that our objective of model quantization has been satisfied with minimal decrease in accuracy. Future iterative developments can improve the accuracy of the quantized model further.

Table IV.
Size of the Base Model M in megabytes before and after Quantization

Base Model size in Mb	41.59658432006836
Quantized Model size in Mb	10.701797485351562

Furthermore, in Table IV, we see that the size of the original base model is around 41.6 Mb whereas its quantized version is

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:33:35 UTC from IEEE Xplore. Restrictions apply.

10.7 Mb. We have successfully reduced the size of the model by 74.28%. This indicates that our model is now small enough to fit in the random access memory (RAM) of many low compute devices. Subsequently we can conclude that our proposed quantized model can be deployed to many modern mobile-set and other types of edge devices.

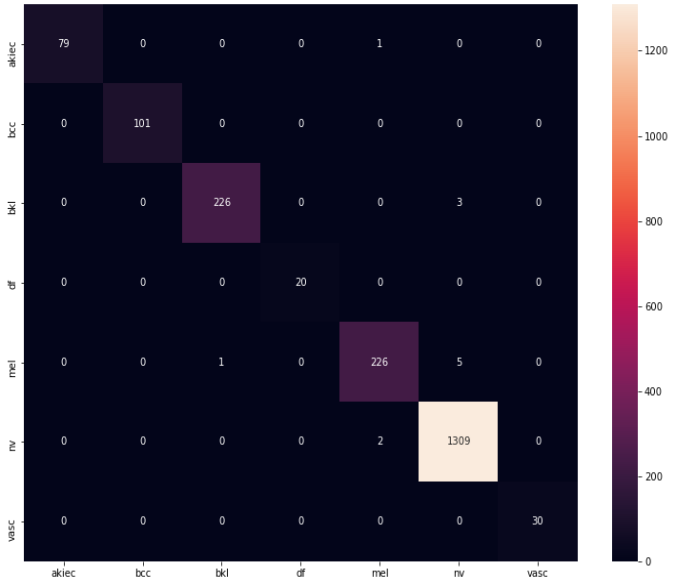


Fig.6. Confusion Matrix for our Proposed Model

The HAM10000 dataset has a great deal of class imbalance. Thus, to obtain a greater sense of clarity of our proposed model, we plot its confusion matrix (Fig.6) using the test dataset. Thus we can see that most of the lesion images have been correctly classified. However, there exists a few erroneously labeled images: for example, in the fifth row, one melanoma lesion has been classified as benign keratosis like lesion, and five melanoma lesions have been misclassified as melanocytic nevi.

V. CONCLUSION AND SCOPE FOR FUTURE WORK

In this paper, we have successfully implemented a low-compute model for classifying skin lesions from dermoscopic images. We use the HAM10000 Dataset for sourcing our training data. In-order to improve the accuracy of our model we normalize the data set and carry out image augmentation prior to training the model. We then carry out the Neural Architecture Search process and obtain the best performing model. We then recreate the architecture of this model separately and train with the augmented image data. Following this, we perform quantization aware training. We quantize layers in the model selectively so as to reach a trade-off between model accuracy, size and latency. Subsequently, we achieve results where the size of our model was reduced by approximately four times with only a minor reduction in accuracy. Furthermore, our base model (non-quantized), achieves better overall and per-class accuracy on the HAM10000 dataset as compared to any other known methodologies, to the best of our knowledge.

In future, work could be continued by deploying and testing such a skin cancer classifier on web applications as well as mobile applications.

REFERENCES

[1] A. G. C. Pacheco and R. R. Krohling, "Recent Advances in Deep Learning Applied To Skin Cancer Detection," *arXiv*, no. NeurIPS, pp. 1–9, 2019.

[2] A. Kwasigroch, M. Grochowski, and A. Mikolajczyk, "Neural architecture search for skin lesion classification," *IEEE Access*, vol. 8, no. January, pp. 9061–9071, 2020, doi: 10.1109/ACCESS.2020.2964424.

[3] I. Maglogiannis and K. K. Delibasis, "Enhancing classification accuracy utilizing globules and dots features in digital dermoscopy," *Comput. Methods Programs Biomed.*, vol. 118, no. 2, pp. 124–133, 2015, doi: 10.1016/j.cmpb.2014.12.001.

[4] C. Barata, M. E. Celebi, and J. S. Marques, "Improving dermoscopy image classification using color constancy," *IEEE J. Biomed. Heal. Informatics*, vol. 19, no. 3, pp. 1146–1152, 2015, doi: 10.1109/JBHI.2014.2336473.

[5] M. Zorman, M. M. Štiglic, P. Kokol, and I. Malčić, "The limitations of decision trees and automatic learning in real world medical decision making," *J. Med. Syst.*, vol. 21, no. 6, pp. 403–415, 1997, doi: 10.1023/A:1022876330390.

[6] D. Ruiz, V. Berenguer, A. Soriano, and B. Sánchez, "A decision support system for the diagnosis of melanoma: A comparative approach," *Expert Syst. Appl.*, vol. 38, no. 12, pp. 15217–15223, 2011, doi: 10.1016/j.eswa.2011.05.079.

[7] B. Erkol, R. H. Moss, R. J. Stanley, W. V Stoecker, and E. Hvatum, "Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes.," *Ski. Res. Technol.*, pp. 17–26, 2005.

[8] S. Gilmore, R. Hofmann-Wellenhof, and H. P. Soyer, "A support vector machine for decision support in melanoma recognition," *Exp. Dermatol.*, vol. 19, no. 9, pp. 830–835, 2010, doi: 10.1111/j.1600-0625.2010.01112.x.

[9] L. Yu, H. Chen, Q. Dou, J. Qin, and P. A. Heng, "Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks," *IEEE Trans. Med. Imaging*, vol. 36, no. 4, pp. 994–1004, 2017, doi: 10.1109/TMI.2016.2642839.

[10] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:33:35 UTC from IEEE Xplore. Restrictions apply.

542, no. 7639, pp. 115–118, 2017, doi: 10.1038/nature21056.

[11] N. C. F. Codella *et al.*, "Deep learning ensembles for melanoma recognition in dermoscopy images," *IBM J. Res. Dev.*, vol. 61, no. 4, pp. 1–28, 2017, doi: 10.1147/JRD.2017.2708299.

[12] N. Gessert *et al.*, "Skin lesion diagnosis using ensembles, unscaled multi-crop evaluation and loss weighting," *arXiv*, pp. 1–5, 2018.

[13] Z. Yu *et al.*, "Melanoma Recognition in Dermoscopy Images via Aggregated Deep Convolutional Features," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 4, pp. 1006–1016, 2019, doi: 10.1109/TBME.2018.2866166.

[14] H. A. Haenssle *et al.*, "Man against Machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists," *Ann. Oncol.*, vol. 29, no. 8, pp. 1836–1842, 2018, doi: 10.1093/annonc/mdy166.

[15] S. S. Han, M. S. Kim, W. Lim, G. H. Park, I. Park, and S. E. Chang, "Classification of the Clinical Images for Benign and Malignant Cutaneous Tumors Using a Deep Learning Algorithm," *J. Invest. Dermatol.*, vol. 138, no. 7, pp. 1529–1538, 2018, doi: 10.1016/j.jid.2018.01.028.

[16] P. Tschandl, C. Rosendahl, and H. Kittler, "Data descriptor: The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Sci. Data*, vol. 5, pp. 1–9, 2018, doi: 10.1038/sdata.2018.161.

[17] H. Jin, Q. Song, and X. Hu, "Auto-keras: An efficient neural architecture search system," *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, pp. 1946–1956, 2019, doi: 10.1145/3292500.3330648.

[18] T. Chen, I. Goodfellow, and J. Shlens, "Net2Net: Accelerating learning via knowledge transfer," *4th Int. Conf. Learn. Represent. ICLR 2016 - Conf. Track Proc.*, pp. 1–12, 2016.

[19] H. Liu, K. Simonyan, and Y. Yang, "Darts: Differentiable architecture search," *arXiv*, pp. 1–13, 2018.

[20] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," *36th Int. Conf. Mach. Learn. ICML 2019*, vol. 2019-June, pp. 10691–10700, 2019.

[21] B. Jacob *et al.*, "Quantization and Training of Neural Networks for Efficient Integer-Arithmetic-Only Inference," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 2704–2713, 2018, doi: 10.1109/CVPR.2018.00286.

[22] P. Carcagnì *et al.*, "Classification of skin lesions by combining multilevel learnings in a DenseNet architecture," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2019, vol. 11751 LNCS, pp. 335–344, doi: 10.1007/978-3-030-30642-7_30.

[23] S. S. Chaturvedi, K. Gupta, and P. S. Prasad, "Skin lesion analyser: An efficient seven-way multi-class skin cancer classification using mobilenet," *arXiv*, pp. 1–11, 2019.

[24] M. Aminur Rab Ratul, M. Hamed Mozaffari, W. S. Lee, and E. Parimbelli, "Skin lesions classification using deep learning based on dilated convolution," *bioRxiv*, 2019, doi: 10.1101/860700.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:33:35 UTC from IEEE Xplore. Restrictions apply.

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 05 | May 2021 www.irjet.net p-ISSN: 2395-0072

Determination of the Probabilities of Landslide Events in Kalimpong, West Bengal

Parvesh Lamba[1], Pooja[2], Amit Dabas[3], Hitesh Khatri[4], Prof. S. Anbu Kumar[5]

[1-5]*(Department of Civil Engineering, Delhi Technological University, Delhi, India.*

ABSTRACT: - *Landslide is more common in mountainous areas due to the low density of rocks, soil, debris etc. The Himalayan region, because of its complex geological features and road offers excellent landslide conditions. It is estimated that 42% of all the most common mudflats in the country fall in the north-east Himalayan region, particularly Darjeeling and Sikkim Himalayas. This leads to the severe loss of life and property which is so early the need to bring in strategies to reduce the impact on the affected areas. Therefore, it is important to understand the relationship between landslides events and rainfall conditions during and before that, primarily in the Himalayan region scenario in Kalimpong. In this paper, a rainfall-based roadmap is proposed and a Bayesian model of soil saturation can be mixed in the Kalimpong region of the Darjeeling Himalayas. Daily rainfall data covering a 10-year period were collected at Teesta's unique rainfall station, Kalimpong (from 2010 to 2019) and landslide (42 total of 146 sites) containing dates to date, corrected from records historical for the period 2010 to 2019. The Bayesian theory for the one-dimensional case was developed. Next, a two-dimensional Bayesian regression model was generated. The probability of occurrence of the landslide is 0.43 if the intensity of rainfall is more than 80 mm/day for the one-dimensional Bayesian formulation and 0.42 for a rainfall period greater than 0.3 days for a range greater than 40 mm / day.*

Author keyword: **Landslide, Kalimpong, Bayesian probability, One dimensional probability, Two-dimensional probability.**

INTRODUCTION

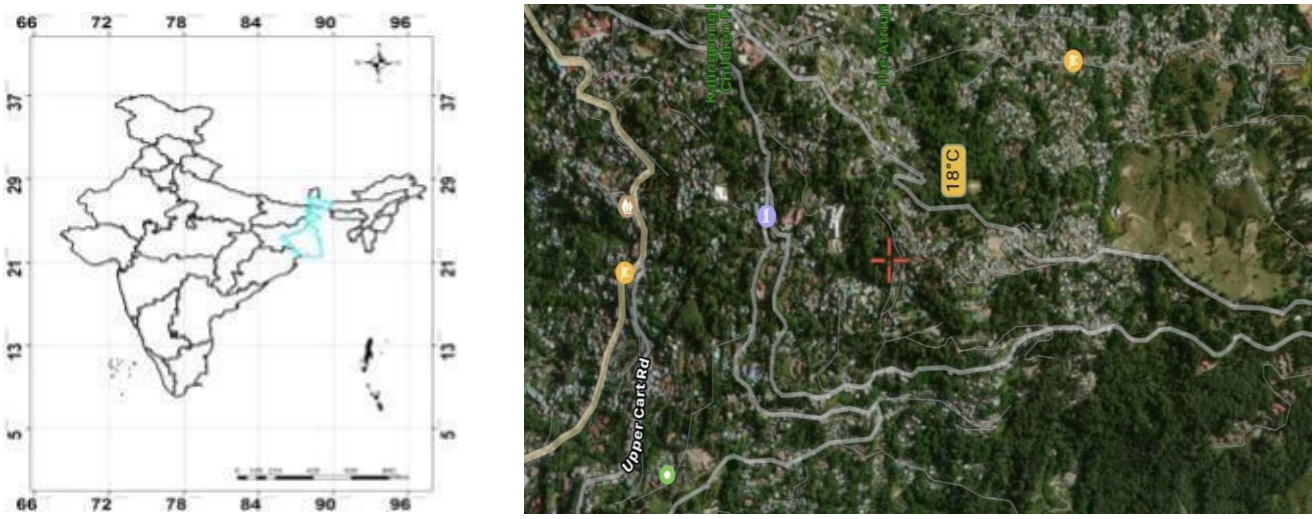
The Indian Himalayan region has been knowingly affected due to the intensification in frequency of landslide incidence. 30% of worldwide landslide events occurs in Himalayan region (GSI Report, 2016) with damage amounting to one billion US$ and loss of 200 people every year. The factors affecting landslides are varied and are usually interconnected. However, the effects due to primary triggering factor are of main concern. As mentioned earlier, three-quarters of landslides in Kalimpong are influenced due to rainfall. The monsoon rainfall in Kalimpong can be defined as low intensity and long duration with disruptions of recurrent heavy gusts which leads to the analysis of shallow landslides. The slightest amount of rainfall required for landslip to befall will define the threshold of the area. Numerous researchers have tried to find a relation between rainfall and initiating of landslips. Few studies have been conducted in the Indian Himalayan region and studied the Darjeeling region; did for North Sikkim and studied a section of the Chamoli region, Uttarakhand. They asserted the importance of finding thresholds using empirical models and emphasized the need to analyse and realize the importance of antecedent rainfall in eliciting landslips. However, some factors like the homogeneity and quality of data may affect the determination of thresholds determined using experimental method. In order to get a well understanding of the effect of rainfall, the analysis of individual rainfall parameters using a probabilistic approach is essential. Moreover, the results obtained from empirical analysis gives a single output which may not always be suitable. Hence, there is a need to move towards a more realistic assessment, by using a probabilistic approach.

STUDY AREA

Kalimpong is a hill station located in the West Bengal region, India. Available at 87.47-89.47 N latitude and 26.07-28.07 E long. The slope of the Western Kalimpong crater has been severely affected by landslides. These sites contain various amounts of interpretations and rubble / rocks of many generations. Global warming is mainly caused by poor physical quality, the erosion of the Teesta River in the toes and hands at short intervals in the event of high rainfall during heavy rainfall. The study area is part of the Darjeeling Himalayan, which includes intratrusted fold thrust belt (FTB) rock from the Precambrian to the Quaternary Ages. soil.

Large amounts of landslides in the area can result in heavy rainfall and untreated water currents or thrusts, which break the bonds, loosen the soil, and break down the particles, leaving the sloping surface area completely saturated. Soil infiltration increases soil fertility thus increasing the surface area. When the soil absorbs too much water, the chances of erosion increase as a result of increased pore pressure. The flow of water causes the rocky weather to collapse along the banks of streams and over time these processes cause the rocks and the other materials to fail leading to landslides. This

area is characterized by a large number of small boundaries known as the Kholas and jhoras. During the monsoon these small streams flow with great force causing constant flooding and depositing a pile of small stones on the rocks. This results in severe backal and headward erosion that results in more severe bank failures in large-scale cracking.



(a) Political Map of India (b)Kalimpong

Fig. 2.1 Location of the Study Area (a) India (b) Kalimpong

CAUSES OF LANDSLIDES IN KALIMPONG

Kalimpong has a history of landslides with the first one recorded in 1899. During the period 2010 to 2019 total of 99 soil erosion events occurred when 61 of the landslides were due to rain.

RAINFALL

Rainfall plays a major role in landslide in this region as heavy rainfall loosens the soil by breaking ties and breaking particles. The magnitude of soil erosion in this region is due to the heavy short rainfall that loosens the soil by breaking the bonds and particle cracks that cause the upper part of the river to become more saturated and causing landslides. Due to heavy rainfall the flat surface is completely filled following erosion. It causes a decrease in the flow of water through the joint and cracks. Rain acts as a softener to soften the cracks. The rain drains the apparatus from the parent's body and expands it under gravity. The percolation of water through the pore space of the soil speeds up the possibility. The water draws down the fall of the earth because it is heavy and can add extra weight to the soil, the rock is overcome by gravity.

The Rainfall database was prepared from the reports of the Geological Survey of India (GSI) and Save the Hills NGO, Kolkata city of Kalimpong during the period 2010 to 2019 as shown in Table 3.1. Fig. 3.1 shows the rainfall data between the months of June and September from 2010 to 2019 based on when it will be analysed.

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
June	317	337	355	248	396	568	327.2	154	184	81
July	656	668	433	424	371	534.4	869.2	811	387.7	821
August	425	525.68	241	401	572	242.3	262.6	432.1	524.5	510.5
September	258	384.1	467	113	265	331.2	366.4	289	346	408.5

Table 3.1: Rainfall (in mm) data for monsoon period in Kalimpong town (2010-2019)

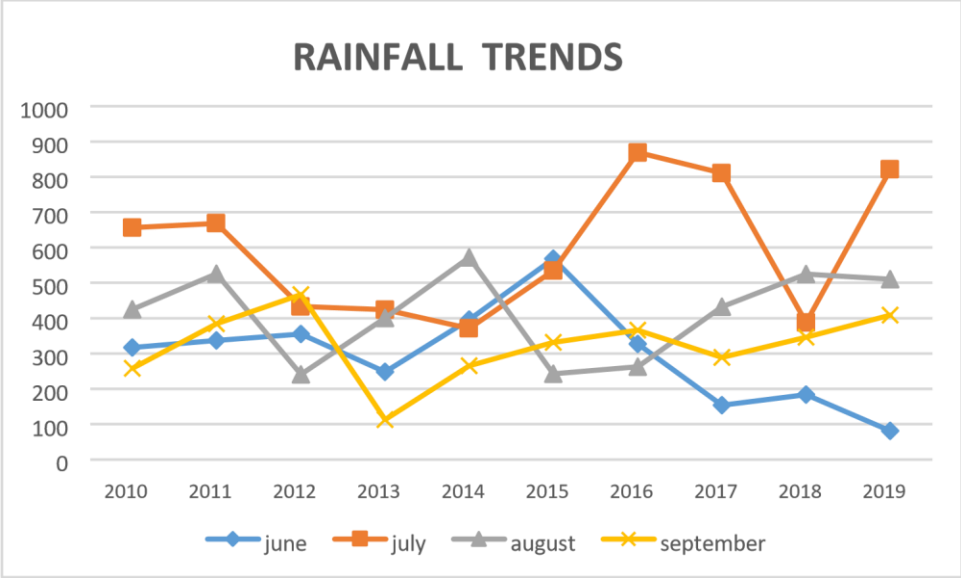


Figure 3.1: Trend in rainfall during the monsoon season

DRAINAGE SYSTEM

The Kalimpong line is thick in topography and is dispersed by low-lying rivals that impact on the Teesta river basin. When the drainage line rises, the frequency of the landfill decreases as the area around the dam network coincides with the distance, so the shear strength of the drainage is reduced around the drainage network.

METHODOLOGY

A probabilistic approach using the Bayesian method is used here to determine the landslide occurrence probability of various precipitation signals. Probability-based methods are advantageous as they integrate different rainfall characteristics or parameters with uncertainties providing a descriptive and qualitative assessment with improved forecasting of the thresholds. Therefore, if its effect is not limited to a particular case of failure or no-failure in a particular rain event, the determination method is inappropriate and a statistical or probabilistic approach may be adopted. It would also ensure that number of landslides N_X is never greater than number of rainfall events N_R, making the probability non-viable, $P(X) = N_X/N_R > 1$.

One-Dimensional Bayesian Analysis

Prior and conditional probabilities are calculated by the application of Bayesian theory which gives the probability of landslide event caused by the various rainfall parameters. The conditional probability is given as P (X|Y) i.e. the probability of landslides due to a particular rainfall parameter (Y) which is found as:

$P(X|Y) = P(Y|X)\,P(X) / P(Y)$

P (Y|X) = Probability of a rainfall event of degree Y given a landslide happens

P (X) = Probability of landslide whether the rainfall event occurs or not,

P (Y) = Probability of rainfall of degree Y whether a landslide happens or not;

For a given N_R, N_X, let the frequency of rainfall events of degree Y be N_Y and the number of rainfall events resulting in landslide be N (Y|X). Now the probability terms can be written in these forms:

$P(X) = N_X/N_R$

$P(Y) = N_Y/N_R$

$P(Y|X) = N(Y|X)/N_X$

Two – Dimensional Bayesian Analysis

Bayesian Probability here estimates the joint conditional probability of two or more influencing parameters.

$$P(X|Y, Z) = P(Y, Z|X). P(X) / P(Y,Z)$$

Here Y, Z denotes the combined probability of any two variables. A pair of rainfall parameters for determining the probability of landslide initiation may be computed using 2-D Bayesian analysis and the significance of this probability is seen when compared with the prior landslide probability values.

ADVANTAGES OF THE METHOD

The major drawback with the deterministic approach is that it only takes into account the rainfall that causes slide initiation i.e. triggering rainfall. Probabilistic (Bayesian) method is advantageous from the fact that it considers all the rainfall events which makes it possible to determine all kinds of uncertainties in terms of probability. A particular rainfall event cannot be said to result in the landslide in such a complex geology every time. In addition, probability calculated using Bayes theorem is dynamic in nature and is easily adaptable to updating processes in future.

LIMITATIONS

Probability computed using Bayes theorem failed when the rainfall event Y did not result in the landslide, meaning that the landslide occurrence probability P (X|Y) should be 0 since P (Y|X) is also 0. This becomes contradicting the real situation. One similar limitation as that of empirical approaches is the accuracy of probabilistic approach with the long-term use of data from the studies in past. Factors such as slope changes, cover on the land, use of the land, rainfall pattern and anthropogenic events affect the occurrence of landslides repeatedly. Bayesian analysis solves this by adding control variables keeping in consideration that probabilistic predictions depend on past data.

APPLICATION OF BAYESIAN APPROACH

One Dimensional Bayesian Probability

Probability of rainfall P (Y) is calculated for N_R = 245 total events of rainfall that took place in study period (2010-2019). The conditional probability P (X|Y) is computed for N_X=42 events of rainfall that initiated landslides.

P (X) = N_X/N_R = 42/245 = 0.17. Now P (Y|X) is calculated considering various rainfall intensities as shown in table 3.2.

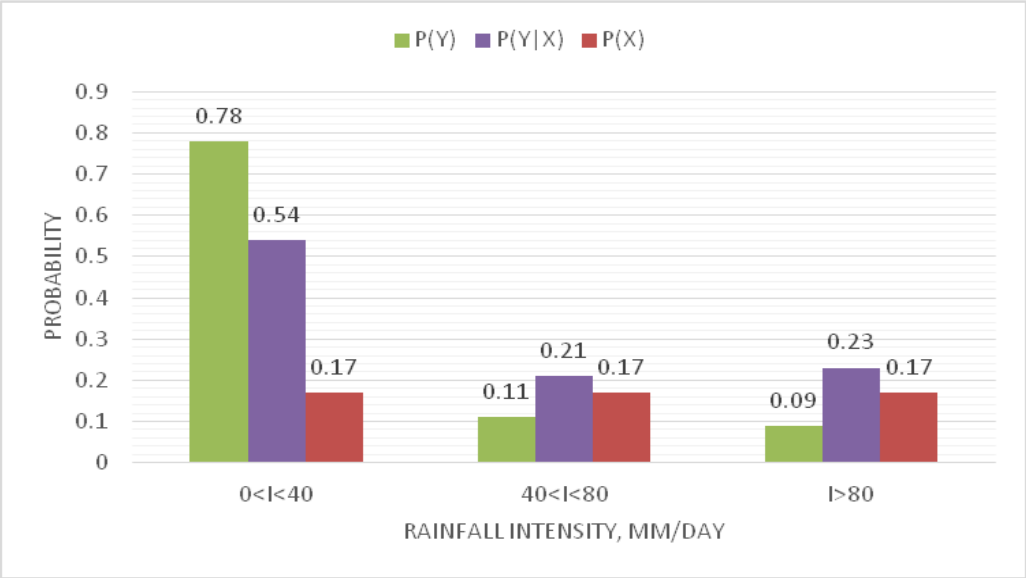
Table 5.1 Rainfall and Landslide Database for 2010-2019

Total rainfall events	Number of daily rainfall intensity with 0<I<40 mm/day	Number of landslide with 0<I<40 mm/day	Number of daily rainfall intensity with 40<I<80 mm/day	Number of landslide with 40<I<80 mm/day	Number of daily rainfall intensity with I>80 mm/day	Number of landslide with I>80 mm/day	Total landslides
245	193	23	29	9	23	10	42

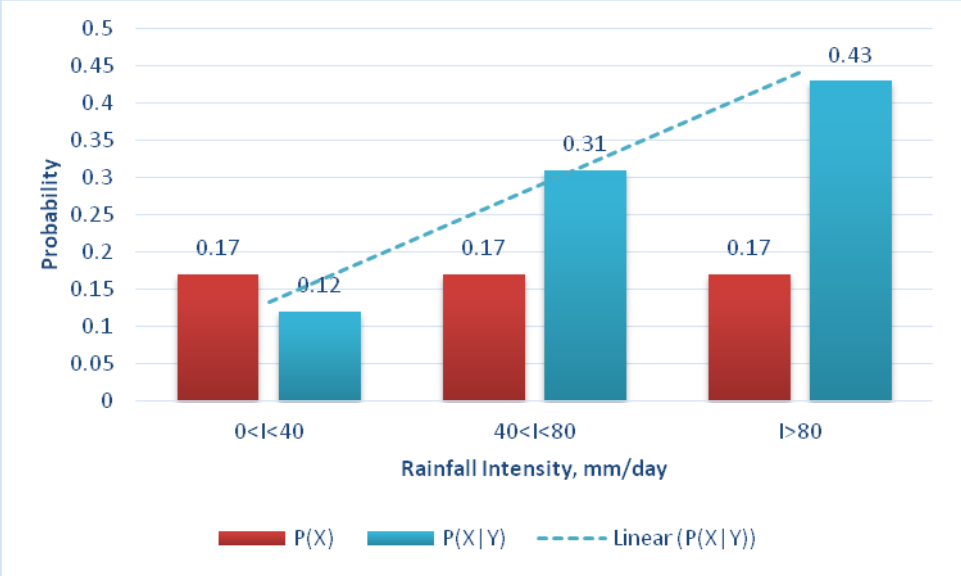
Table 5.2 One-Dimensional analysis for the study period (2010-2019)

Rainfall Intensity Classification	Rainfall events	P(X)	P(Y)	P(Y|X)	P(X|Y)	Total Landslides	Total rainfall events
0<I<40	193	0.17	0.78	0.54	0.12	42	245

40<I<80	29	0.17	0.11	0.21	0.31
I>80	23	0.17	0.09	0.23	0.43



(a)



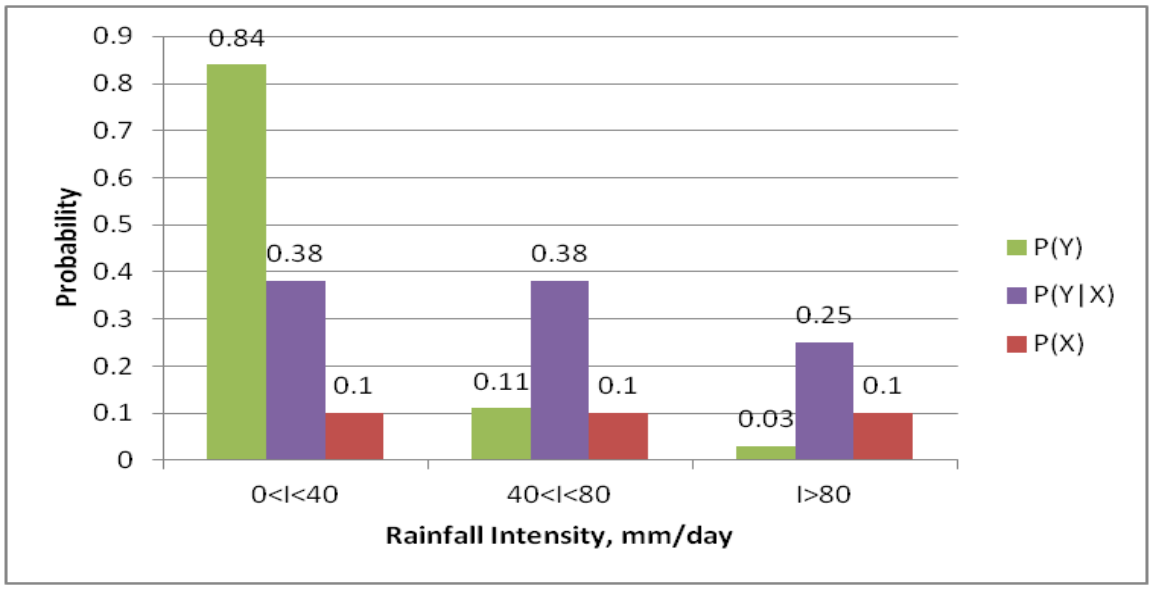
(b)

Table 5.3 Rainfall and Landslide Database for the monsoon period of the year 2019 separately to calculate 2D probabilities.

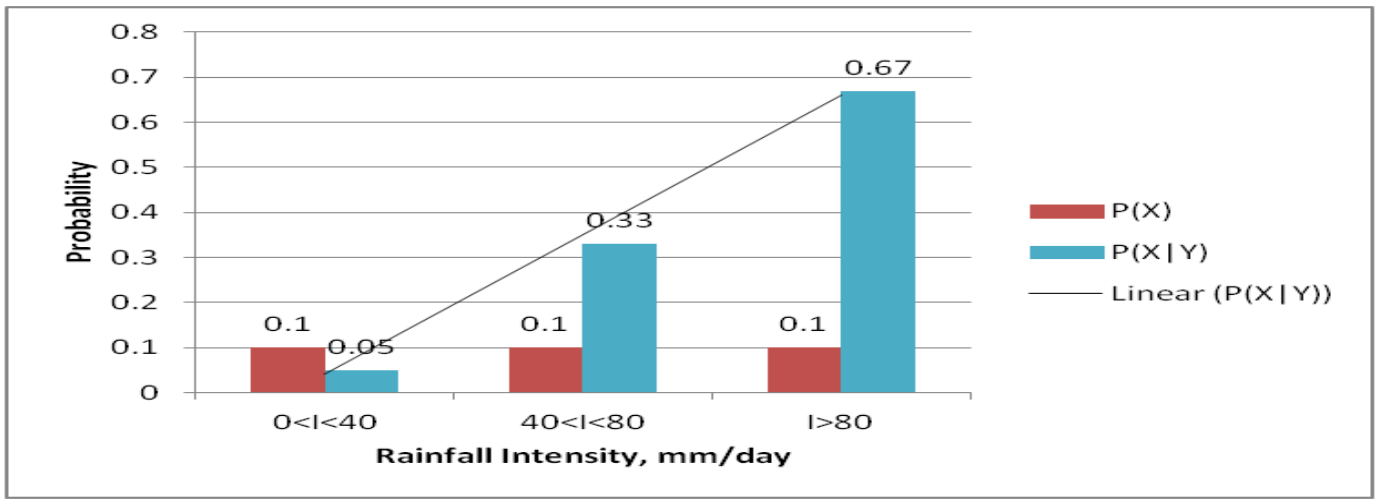
Total rainfall events	Number of daily rainfall intensity with 0<I<40 mm/day	Number of landslide with 0<I<40 mm/day	Number of daily rainfall intensity with 40<I<80 mm/day	Number of landslide with 40<I<80 mm/day	Number of daily rainfall intensity with I>80 mm/day	Number of landslide with I>80 mm/day	Total landslides
77	65	3	9	3	3	2	8

Table 5.4 One-Dimensional analysis for the monsoon period of the year 2019 separately to calculate 2-D probabilities.

Rainfall intensity classification	Rainfall Events	P(X)	P(Y)	P(Y|X)	P(X|Y)	Total Landslide	Total rainfall events
0<I<40	65	0.1	0.84	0.38	0.05	8	77
40<I<80	9	0.1	0.11	0.38	0.33		
I>80	3	0.1	0.03	0.25	0.67		



(c)



(d)

Two-Dimensional Bayesian Probability

To calculate the probabilities in 2-D, the Bayesian probability model requires more detailed data collection with the intensity and duration of the individual rainfall event. Since the hourly or weekly rainfall statistics from the past were not readily available, we were only able to obtain the required data for 2019 (June-July).

Table 5.3 provides the required data for the calculation while table 5.4 shows the probability computations.

Four different combinations of intensity-duration are taken for the computations.

Table 5.3 Rainfall intensity (mm/day)-duration (days) database for Monsoon 2019

Rainfall intensity-duration	RID resulting in landslide	RID not resulting in landslide	Total RID	Total Landslide s	Total RID Events
I<=40 & D<=0.3	2	41	43	8	77
I<=40 & D>0.3	1	21	22		
I>40 & D<=0.3	2	3	5		
I>40 & D>0.3	3	4	7		

Table 5.4 Probability computations.

Rainfall intensity Duration (RID)	P (I|D)	P (I,D|X)	P (X)	P (X|I,D)
I<=40 & D<=0.3	0.56	0.25	0.1	0.04
I<=40 & D>0.3	0.29	0.13	0.1	0.04
I>40 & D<=0.3	0.06	0.25	0.1	0.42

I>40 & D>0.3	0.09	0.38	0.1	0.42

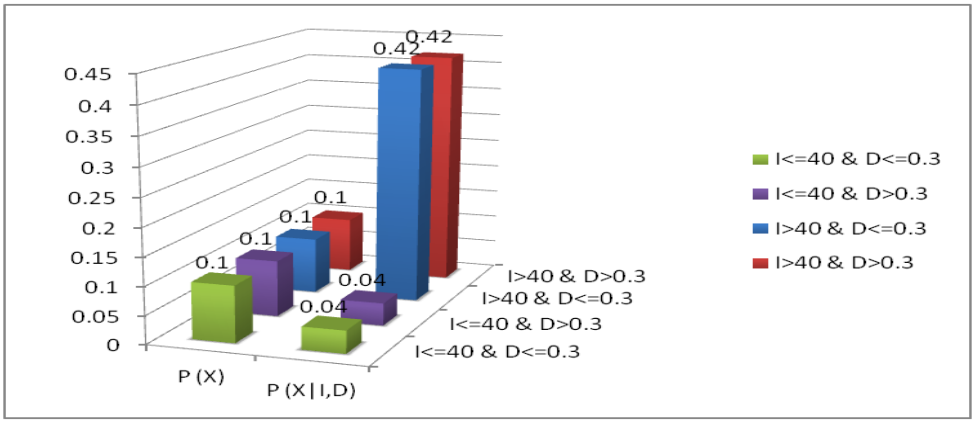


Fig. 5.3 Histogram showing landslide probability with joint parameters

Fig. 5.3 shows the graphical representation of the conditional probabilities. The significance of the parameters considered in the computations can be clearly seen from the difference in the values of probability.

DISCUSSIONS

There are a number of habitat patterns that correlate the difficulty limits with the prevalence of landslides. Such an approach is not always useful in determining the outcome and fully assessing landslides.

The analysis was applied over a ten-year period (2010–2019), with multiple latency parameters such as length, intensity and event degradation. This can often be improved after adding some basic and additional variables and depth degradation and landslide knowledge of a type of analysis often performed for a mountain range area. Analysis has been made of 2 different variants of Thomas Bayes' theory; One-dimensional and two-dimensional. One-dimensional calculation of landslide prevalence is based entirely on the probability of any trough parameter, whereas 2-dimensional calculates the probability of landslide propagation for any 2 trough parameter.

To verify the results, sensitivity analysis was performed by variableizing the input parameters and identifying the corresponding correction by chance. The results showed the inclination of the associative degree to change the severity of the fall, and so the landslides were shifted to a certain number of combinations of landslides. However, there is little choice with modifying rainfall opportunities. Comparative analysis has been done for 2-D possibility; therefore, the results are extra sensitive compared to 1-D results. Investigations showed that the intensity and incidence decrease were very sensitive when compared with other combinations.

CONCLUSIONS

The results indicate that of the rainfall parameters used in the study (depth of rainfall, rainfall duration and excess rainfall), the impact of rainfall intensity is very important in each case. The probability of erosion is 0.43 of rainfall greater than 80 mm / day in 1-D case. However, the use of 2-D approach gives a better picture of the landslide events when compared to the 1-D case. The probabilities reach to a maximum value of 0.42 for the rainfall with duration greater than 0.3 days having intensity of rainfall higher than 40 mm/day for the two dimensional landslide cases. The results also show that its potential depends on the precision and accuracy of the data recorded. Samples having limited amount data are of limited knowledge and several variations in the geology of the region may indicate conflicting possibilities. The use of an over-the-top prediction method is one of the best ways to develop an early warning system for the affected regions when explaining using threshold models.

From the present research the following conclusions are drawn:

The probability of occurrence of the landslide is 0.43 if the intensity of rainfall is more than 80 mm/day for the one dimensional Bayesian formulation and 0.42 for a rainfall period greater than 0.3 days for a range greater than 40 mm / day.

ACKNOWLEDGEMENTS: - The authors would like to thank the Delhi Technological University staff, SAVE THE HILL NGO for the support provided to carry out the research.

REFERENCES

1. S. Acharyya. A note on the geology of the eastern himalayan foot-hills of darjeeling-western duars area. Quart. Jour. Geol. Min. Met. Soc. Ind, 40:43–44, 1968. https://shodhganga.inflibnet.ac.in/bitstream/10603/165221/18/18_bibliography.pdf
2. F. Agterberg, G. F. Bonham-Carter, and D. Wright. Statistical pattern integration for mineral exploration. In Computer applications in resource estimation, pages 1–21. Elsevier, 1990. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.138.3094&rep=rep1&type=pdf
3. P. Aleotti. A warning system for rainfall-induced shallow failures. Engineering Geology, 73(3-4): 247– 265, 2004. https://www.researchgate.net/publication/223232501_A_warning_system_for_rainfallinduced_shallow_failures
4. M. Bacchini and A. Zannoni. Relations between rainfall and triggering of debris-flow: case study of cancia (dolomites, northeastern italy). Natural Hazards and Earth System Science, 3(1/2):71– 79, 2003. https://www.nat-hazards-earth-syst-sci.net/3/71/2003/
5. S. Barbero, D. Rabuffetti, and M. Zaccagnino. Una metodologia per la definizione delle soglie pluviometriche a supporto dellemissione dellallertamento. Proc. 29th Convegno Nazionale di Idraulica e Costruzioni Idrauliche, Trento, pages 7–10, 2004. http://www.idrologia.polito.it/web2/persone/team- members/claps/pubblicazioni/
6. M. A. Bean. Probability: the science of uncertainty with applications to investments, insurance, and engineering, volume 6. American Mathematical Soc., 2009. https://bookstore.ams.org/amstext-6/
7. M. Berti, M. Martina, S. Franceschini, S. Pignone, A. Simoni, and M. Pizziolo. Probabilistic rainfall thresholds for landslide occurrence using a bayesian approach. Journal of Geophysical Research: Earth Surface, 117(F4), 2012. https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2012JF002367
8. C. Bonnard and F. Noverraz. Influence of climate change on large landslides: Assessment of long- term movements and trends. In International Conference on Landslides: Causes, Impacts and Countermeasures, number CONF, pages 121–138. VGE, 2001. https://infoscience.epfl.ch/record/94492?of=HB
9. N. Caine. The rainfall intensity-duration control of shallow landslides and debris flows. Ge- ografiska annaler: series A, physical geography, 62(1-2):23–27, 1980. https://www.jstor.org/stable/520449?seq=1
10. R. H. Campbell. Soil slip, debris flows, and rainstorms in the santa monica mountains and vicinity, southern california. US Geological survey professional paper, 851:51, 1975. https://pubs.usgs.gov/pp/0851/report.pdf

Authors Profile



Parvesh He is currently pursuing his bachelor of technology in civil engineering from Delhi Technological University, Delhi.



Pooja She is currently pursuing his bachelor of technology in civil engineering from Delhi Technological University, Delhi.



Hitesh Khatri He is currently pursuing his bachelor of technology in civil engineering from Delhi Technological University, Delhi.



Amit Dabas He is currently pursuing his bachelor of technology in civil engineering from Delhi Technological University, Delhi.



Dr S. ANBUKUMAR He is currently working as a associate professor in department of Civil Engineering in Delhi Technological University, Delhi. Completed his masters of technology from IIT Madaras, Chennai and received his PhD from Delhi Technological University.

Silicon
https://doi.org/10.1007/s12633-021-01191-4

ORIGINAL PAPER

Dielectric Engineered Schottky Barrier MOSFET for Biosensor Applications: Proposal and Investigation

Rahul Singh[1] · Shweta Kaim[1] · Rani MedhaShree[1] · Anil Kumar[1] · Sumit Kale[1]

Received: 22 March 2021 / Accepted: 1 June 2021
© Springer Nature B.V. 2021

Abstract

In this paper, for the first time, we have investigated a Dielectric Engineered Schottky Barrier MOSFET (DE-SBMOS) for Biosensor applications. The DE SBMOS uses dielectric engineering and modulation approach to detect biomolecules. The DE SBMOS incorporated the high-k gate dielectric at the source side and low-k gate dielectric at the drain side. In addition, a nano gap cavity is created near the source region by etching the gate metal precisely. The presence of biomolecules in the cavity results in change in drive current of the device. The sensitivity of the proposed device is measured in terms of threshold voltage (V_{TH}), I_{ON} current, I_{OFF} current, Subthreshold Swing (SS), and I_{DS} current. In later part, misalignment effect of cavity length with high-k dielectric material length examined and that result in three different modes i.e. Underlapped Biosensor (UBS), Aligned Biosensor (ABS), Overlapped Biosensor (OBS). Based on aforementioned modes, the proposed device shows the optimal choice for OBS mode.

Keywords Dielectric engineering · Schottky barrier · Biosensors · UBS · ABS · OBS · Nano gap · Cavity · Biomolecule

1 Introduction

Over the last few years, biosensor attracts the researchers in the field of clinical diagnosis, food and veterinary technologies, agriculture industry, industrial processing, medical diagnosis. An early-stage detection and diagnosis of the diseases is very difficult such as cancer, HIV, diabetes, etc. [1]. The conventional label detection method does not provide the high efficiency because this method changes the inherent property of the biomolecules. Meanwhile, label free biosensing method has gained attraction from researchers. FET based biosensor device has always been attracting researchers due to the numerous advantages: 1) Simple design 2) Low cost to synthesize 3) Provides fast detection because of high specificity and affinity binding with target biomolecules. 4) Label-free detection 5) Compatibility with traditional CMOS process flow. 6) Successful on-chip integration 7) Low response time and high sensitivity for the detection of biomolecules etc. [3–8]. For the diagnosis of various diseases many biosensor-based devices have been widely studied. However, Schottky Barrier MOSFETs (SBMOS) have attracted significant interest in recent times due to their future potential for high-performance switching, amplification, mixed-mode and analog RF applications [10–16]. The biological properties of biomolecules are converted into analogous electrical properties in biosensor technology. This concept has been used in the proposed device by varying the dielectric properties the electrical properties of the device will change. Based on variation in electrical properties, the sensitivity of the device has been obtained. Various detection methods have been used to detect the presence or absence of biomolecules, but the most efficient method is Dielectric Modulation (DM) method. This approach is effective for both charged and neutral biomolecules [2, 3]. In DM method, the values of dielectric material are incremented and the variation in electrical parameter is sensed. To use a device as a biosensor, a nano cavity is created to capture bio-molecules. In the proposed device it is used in gate 2 metal electrode [2, 6–9] to detect the various biomolecules different dielectric constant values are selected accordingly [17–22]. The dielectric constant of biomolecules ranging from 2 to 5 is categorized as low-k molecules (Uricase = 1.54, Streptavidin = 2.1, protein = 2.50, Biotin = 2.63, 3-aminopropyltriethoxysilane = 3.57, etc.) and the dielectric constant of biomolecules ranging from 10 to 12 is categorized as high-k molecules (Gelatin, etc.) [24–27]. Sensitivity parameters for different biomolecules using proposed device, namely, V_{TH} sensitivity, SS sensitivity, I_{ON}

✉ Sumit Kale
sumitkale785@gmail.com

[1] Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India

Published online: 09 June 2021

Table 1 Typical parameter values of Proposed DE-SBMOS with dielectric Modulation

Parameters	Values
Channel Length, L	50 nm
HfO_2 gate oxide thickness	2 nm
SiO_2 gate oxide thickness	2 nm
Si Channel thickness	30 nm
Channel doping, N_A	$1x10^{15}cm^{-3}$
Length of Source/Drain regions	25 nm
Gate 1 thickness	2 nm
Gate 2 thickness	15 nm
Work function gate 1	4.4 eV
Work function gate 2	4.8 eV
Electron SB Height	0.65 eV
Hole SB Height	0.47 eV
Effective mass of electron	0.32 m_0
Effective mass of hole	0.54 m_0

sensitivity, I_{OFF} sensitivity, etc., having L = 50 nm, has considered for examination. The variation of sensitivity parameters has been carried out by changing the length of cavity and the high-k material [1].

2 Device Structure and Simulation Setup

Figure 1 shows the 2-D schematic of our proposed DE-SBMOS biosensor device. Table 1, shows typical parameters values taken while simulating the device structure. Schottky barrier MOSFET uses metal contacts so that Schottky barrier height decreases that will eventually reduce leakage current. Most common contact metals for Schottky MOSFETs are Silicon Carbide, Silicides such as Cobalt Disilicide, Nickel Silicide, and Platinum Silicides which offers low resistance and lower leakage current. In the proposed device, Nickel Silicide has been used as Source/Drain material to form Schottky Barrier MOSFET (SBMOS). Nickel Silicides are generally thermally and chemically stable. Nickel Silicide has low electrical resistivity. The length of Source/Drain region is 25 nm. In the structure of DE-SBMOS device L_{ch} is the sum of L_{HfO2} and L_{SiO2}, where L_{HfO2} represents the length of high-k dielectric gate material (HfO_2) and L_{SiO2} represents the length of low-k dielectric material (SiO_2). NiSi is employed as source/drain contact [11, 13–16]. For gate 1, chromium (Cr) metal has been used, having work function around 4.4 eV, meanwhile for gate 2, silver(Ag) has been used (work function around 4.8). Silver has high affinity for a variety of biomolecules and also have high conductivity. Chromium has good chemical bonding properties towards silver. So, silver-chromium pair provide strength to the structure [6, 22, 23].

To perform the misalignment experiment, three modes have been used. Underlapped biosensor (UBS), in which the ratio of the L_{Cavity} to $L_{High\text{-}k}$ has been kept at 4:5,and in overlapped biosensor mode(OBS) L_{Cavity} to $L_{High\text{-}k}$ has been kept at 6:5. In the third mode, that is aligned biosensor mode(ABS), the ratio is 1:1.

Simulation have been done using SILVACO ATLAS 2-D numerical device simulator [29]. The Universal Schottky Tunneling (UST) model has been included to take in consideration Schottky tunneling current. This model assumes a triangular barrier and uses Wentzel–Kramers–Brillouin approximation. Also, we have included thermionic emission models for the estimation of thermionic emission currents. In addition, Shockley-Read-Hall (SRH) recombination is used for fixed carrier lifetimes and Lombardi mobility models which includes the temperature, doping and field dependent effects, inversion-layer Lombardi constant voltage and temperature (CVT) mobility model takes into account the effect of transverse fields, along with doping and temperature dependent

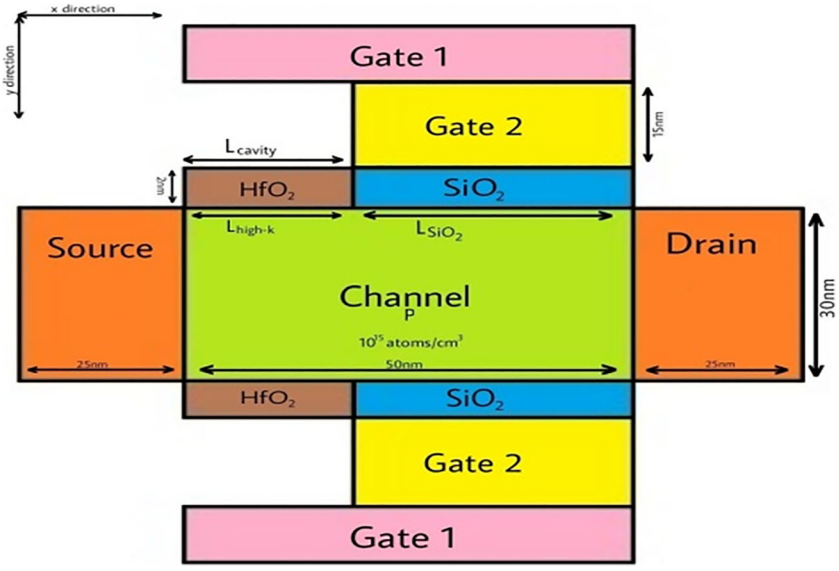


Fig. 1 Structure of the proposed DE-SBMOS with dielectric modulation

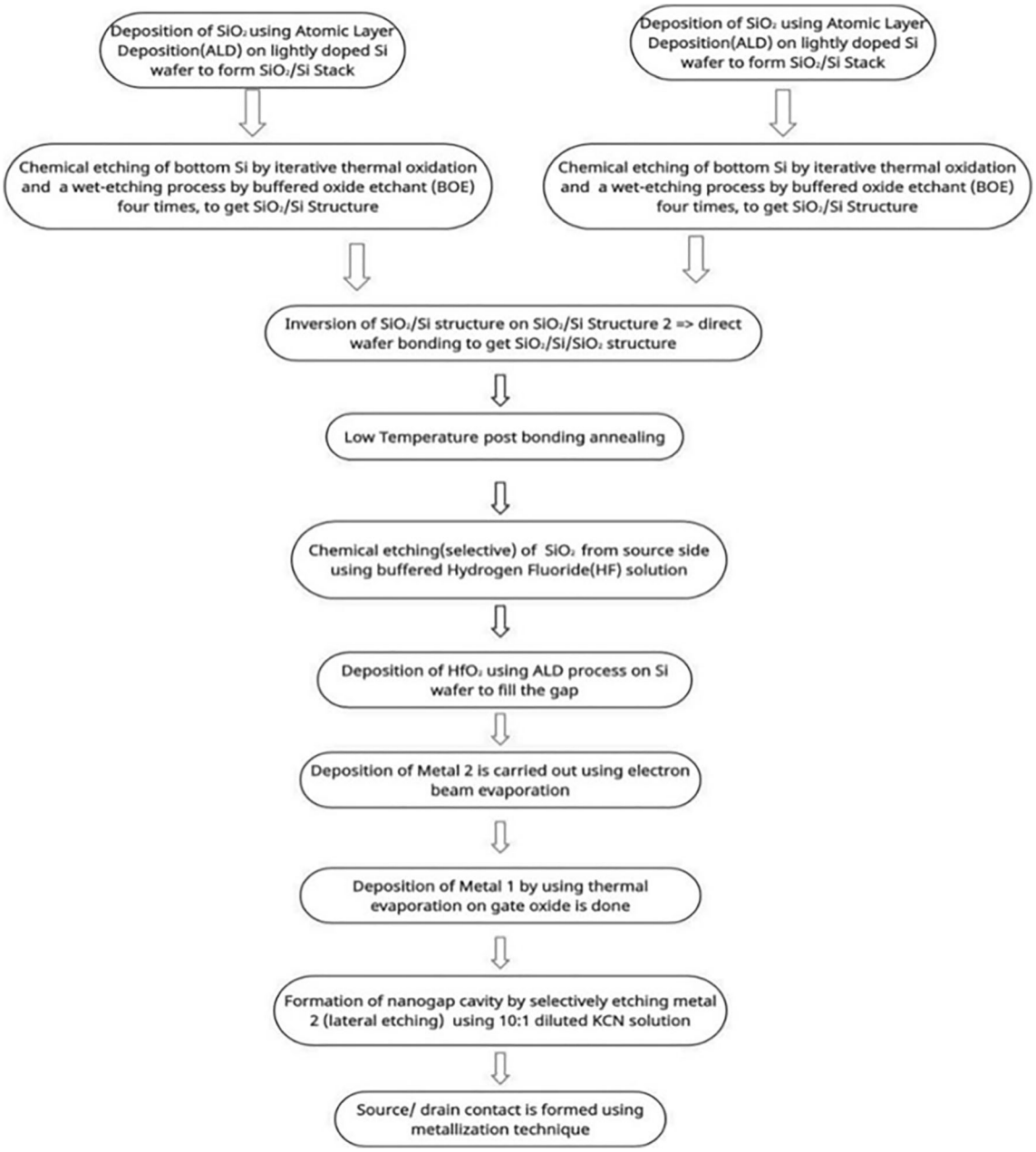
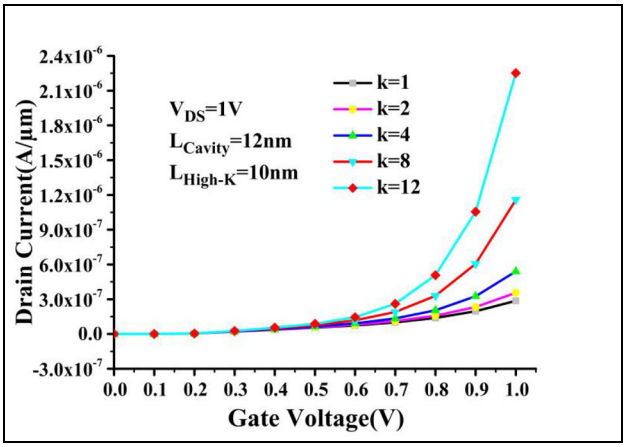


Fig. 2 Possible fabrication flow the proposed DE-SBMOS

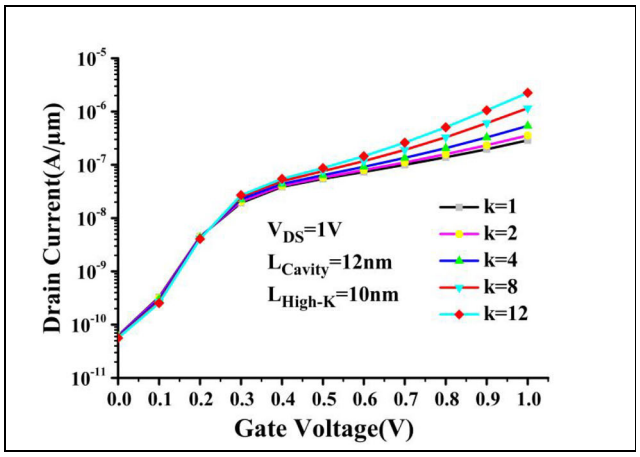
parts of the mobility and the Shockley–Read–Hall (SRH) model simulates the leakage currents that exist due to thermal generation. Furthermore, BGN model is also used to account tunneling effect more accurately in heavily doped region [28, 29]. Richardson's constants for electrons and holes are taken as 30 and 110 A/Kcm^2.

3 Proposed Fabrication Steps of DE-SBMOS Based Biosensor

Figure 2 shows the possible fabrication steps of the proposed DE-SBMOS for Biosensor applications [30–34]. The lightly doped (1×10^{15} cm^{-3}) p-type silicon wafer can be considered as a starting material. The formation of SiO_2 is done for both top and bottom surface using Atomic Layer Deposition (ALD) on silicon wafer [11]. Subsequently, the selective etching of SiO_2 from source side using buffered Hydrogen Fluoride (HF) solution has been carried out [11]. Next, the deposition of HfO_2 gate oxide layer over the uncovered silicon wafer near the source/channel interface is performed using Atomic layer deposition process [31–33]. After that, deposition of Metal 2 (Silver) is carried out using electron beam evaporation followed by deposition of Metal 1 (Chromium) by using thermal evaporation on gate oxide is carried out. For



(a)



(b)

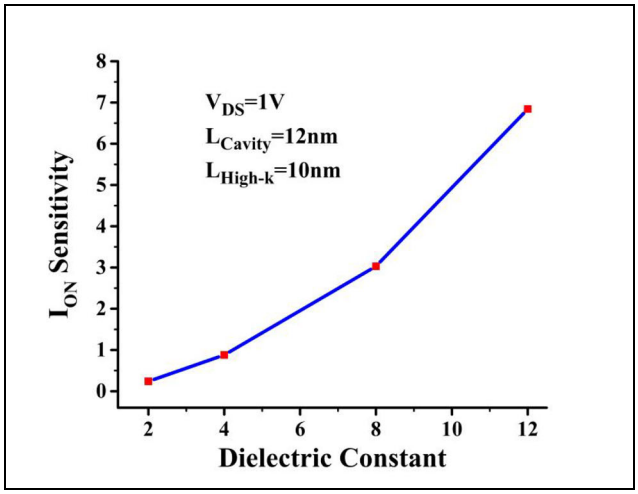
Fig. 3 (a) Linear Transfer characteristics (b) Logarithmic I_D-V_{GS} of DE-SBMOS based biosensor at different dielectric constants (k = 1, 2, 4, 8, & 12)

sensing the biomolecules, the metal 2 (Silver) is selectively etched out (lateral etching) to form the nano gap cavity using 10:1 diluted KCN solution [1, 11]. Next, as nickel metal can be deposited by electron-beam evaporation. Rapid thermal annealing can be carried to form NiSi interface. Finally, the contacts can be formed [35–45].

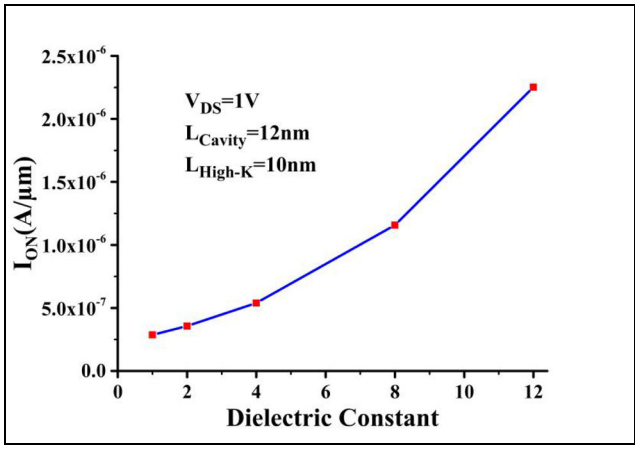
4 Result and Discussion

4.1 Sensitivity Parameters of DE-SBMOS Based Biosensor Device for OBS Mode

In Fig. 3, the length of cavity and high-k material is kept at 12 nm and10nm, respectively, with $V_{DS} = 1$ V. For the increase in dielectric constant of material under the gate electrode as 1, 2, 4, 8, and 12, resulting the drive current of the device varies with the applied voltage. This is due to



(a)



(b)

Fig. 4 (a) I_{ON} Sensitivity of DE-SBMOS based biosensor at different dielectric constants (k = 2,4,8,&12) and (b) I_{ON} Current of DE-SBMOS based biosensor at different dielectric constants (k = 2, 4, 8, & 12)

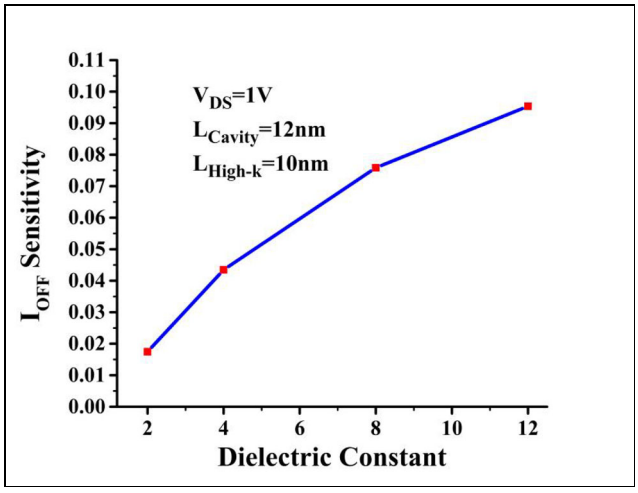
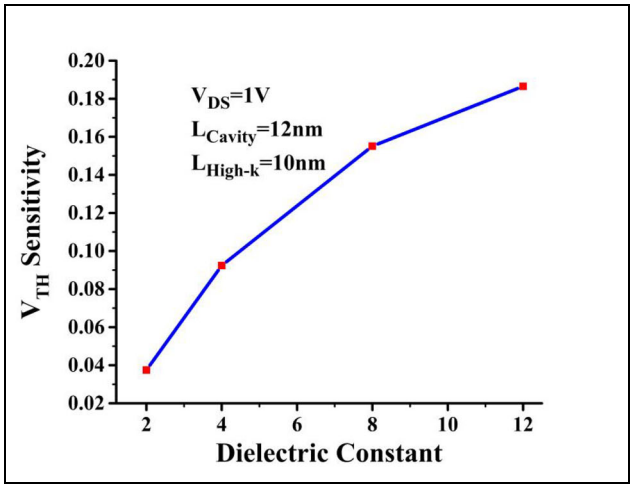
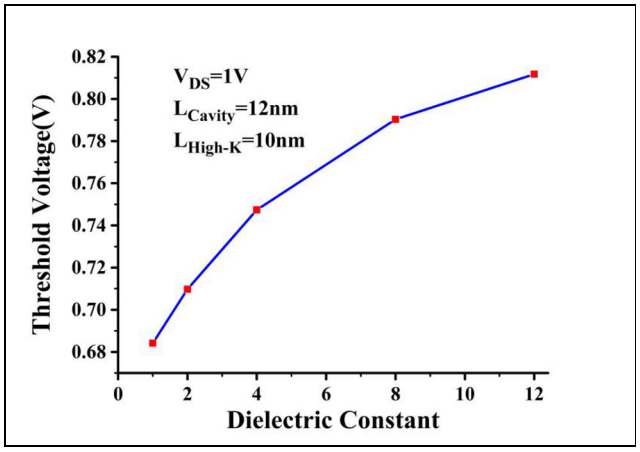


Fig. 5 I_{OFF} Sensitivity of DE-SBMOS based biosensor at different dielectric constants (k = 2,4,8, &12)

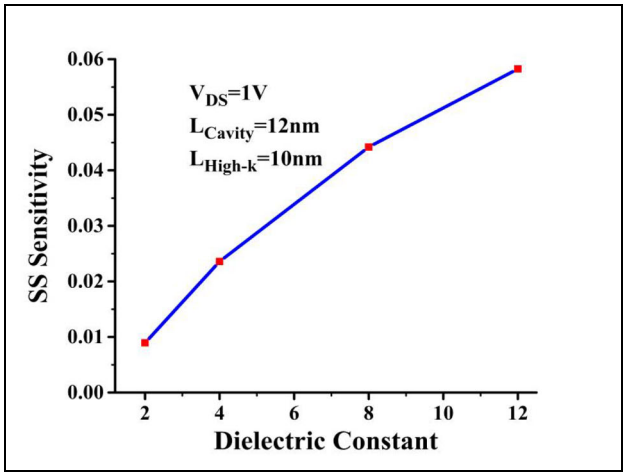


(a)

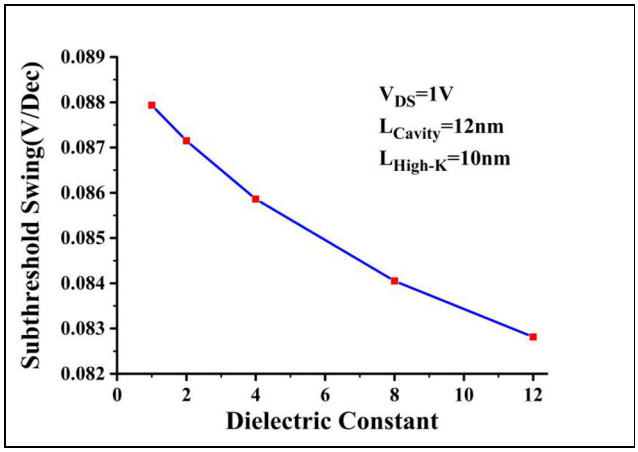


(b)

Fig. 6 (a) V_{TH} Sensitivity of DE-SBMOS based biosensor at different dielectric constants (k = 2,4,8, &12) and (b) Threshold Voltage (V_{TH}) of DE-SBMOS at different dielectric constants (k = 2, 4, 8, & 12)



(a)



(b)

Fig. 7 SS Sensitivity of DE-SBMOS based biosensor at different dielectric constants (k = 2,4,8, &12)

decrease in the Schottky barrier width of the device at the source end. Because of high-k dielectric material near source side, more carriers will flow from the source/channel interface region towards the drain side. Therefore, significant variation in drive current is observed for different dielectric constant of material. The use of low–k dielectric material at the drain side contributes in reduction of ambipolar current. As a result, the proposed device exhibits low ambipolar current due to reduction in electric field supports less band bending.

The performance of biosensor has been observed with the help of analysis of sensitivity. Sensitivity is used to detect the presence/absence of target. The sensitivity of the device is checked by examine the change in the current value in presence of various biomolecules to the current value in absence of the biomolecules in the cavity. The expression for the sensitivity is defined as:

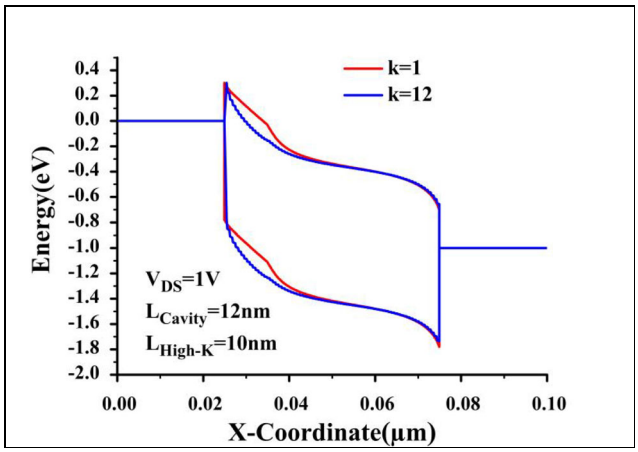


Fig. 8 Plots of the energy band profile for k = 1 and k = 12, DE-SBMOS at V_{DS} = 1 V and V_{GS} = 1 V

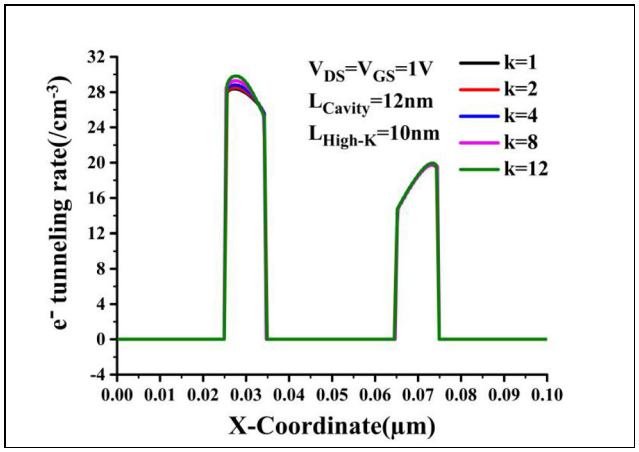


Fig. 9 Tunneling rate of DE-SBMOS based biosensor at different dielectric constants (k = 1, 2, 4, 8, & 12)

$$S = \left| \frac{P_{bio}}{P_{air}} - 1 \right|$$

In the expression (1), S is the sensitivity parameter, P_{bio} is the electrical parameter value in presence of biomolecules in the cavity and P_{air} is the electrical parameter value in absence of biomolecules in the cavity.

The drive current increases with the increment in dielectric constant of material, as electric field is higher at the source/channel junction which results in a large value of current because of low schottky barrier height. The same effect is seen with the I_{ON} sensitivity, it increases with increase in dielectric constant value. The I_{ON} current must be as high as possible for high sensitivity. The same is depicted from Fig. 4.

In Fig. 5, I_{OFF} sensitivity of DE-SBMOS increases with the increment in the dielectric constant value and off current decreases with increase in dielectric constant value because at drain/channel junction, electric field is low and barrier height increases therefore a small amount of I_{off} current. The I_{OFF} current must be as low as possible in order to have higher sensitivity. The same is depicted from Fig. 5, as dielectric constant value increases the I_{OFF} decreases and sensitivity increases.

In Fig. 6, V_{TH} sensitivity of DE-SBMOS for different value of dielectric constant has been shown at optimized cavity length of 12 nm and high-k material length of 10 nm. The V_{TH} sensitivity and threshold voltage both increases with increment in the dielectric constant. As dielectric constant increases, overlap capacitance and junction capacitance increases that results in higher threshold value. From Fig. 4, it is clear that as V_{TH} increases I_{OFF} current decreases which results in a lower leakage current. The Subthreshold Slope (Subthreshold Swing) (SS) is the major parameter for calculating the off-state current. Furthermore, SS is calculated as

$$ss = \frac{1}{\frac{\partial(log(I_d))}{\partial V_{gs}}}$$

The SS value is extracted by calculating the inverse of maximum slope of V_{gs} versus $\log(I_d)$ curve.

In Fig. 7, as the dielectric constant value incremented, the SS value decreases and SS sensitivity will increase which is result of high electrostatic coupling between gate and source/channel region. When there is higher threshold voltage there will be less subthreshold leakage current and low subtreshold swing. From above Fig. 7, we conclude that for higher sensitivity the SS value must be as low as possible.

The energy band profile with respect to change in k for DE-SBMOS is illustrated in Fig. 8. The band banding of valance and conduction band for k = 1 and k = 12 in the cavity region is shown. This change in energy band profile essentially decreases the schottky barrier width at the source/channel

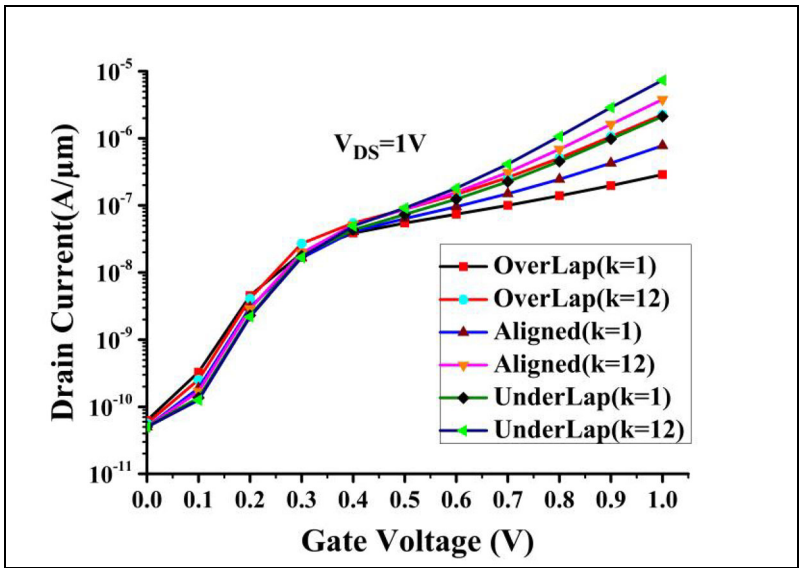


Fig. 10 Transfer characteristics of DE-SBMOS based biosensor for UBS, OBS & ABS mode (V_{DS} = 1 V, k = 1&12)

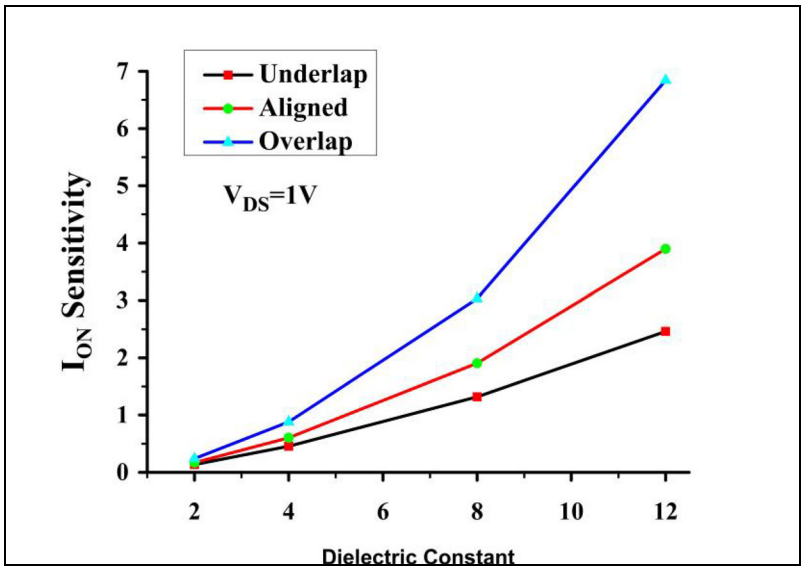


Fig. 11 I_{ON} Sensitivity of DE-SBMOS based biosensor for UBS, OBS & ABS mode (V_{DS} = 1 V, k = 2, 4, 8, & 12)

junction and increases the transportation of charge carrier from source to drain.

In Fig. 9, there is an increase in the electron tunneling rate with the increment in high-k dielectric constant value due to decrease in the schottky barrier width at the source/channel junction. As a result, large number of charge carriers can tunnel from source side towards drain end.

4.2 Sensitivity Parameters of DE-SBMOS Based Biosensor Device in the UBS, OBS, and ABS Modes

In Fig. 10, the drain current increases with the increment in dielectric value. This is due to decrease in schottky barrier width at the source/channel junction. The UBS mode has the highest value among all three modes.

From Fig. 11, it can be seen that the I_{ON} sensitivity increases with the increment in the dielectric value. The OBS mode has the highest I_{ON} sensitivity as compare to all other modes. This mode is the optimal mode and it is used throughout the paper.

From Fig. 12, it is seen that the I_{OFF} Sensitivity increases with the increment in the dielectric value. The OBS mode has the highest I_{OFF} sensitivity as compare to all other modes.

From Fig. 13, it is seen that the ABS and UBS modes are less sensitive to the variation of V_{TH} change as compare to OBS mode.

From Fig. 14, SS Sensitivity increases with the increment in the dielectric value. The OBS mode has the high SS sensitivity as compare to rest of the modes.

The Table 2 reflects the sensitivity parameters for UBS, ABS, & OBS modes. The OBS mode shows the high sensitivity towards I_{DS} Sensitivity, V_{TH} Sensitivity, I_{ON} Sensitivity as compare to other two modes. The length of cavity is varied with respect to high-k materials at V_{DS} = 1 V. Hence, from the table, OBS mode is the optimized mode for DE-SBMOS based Biosensor Device.

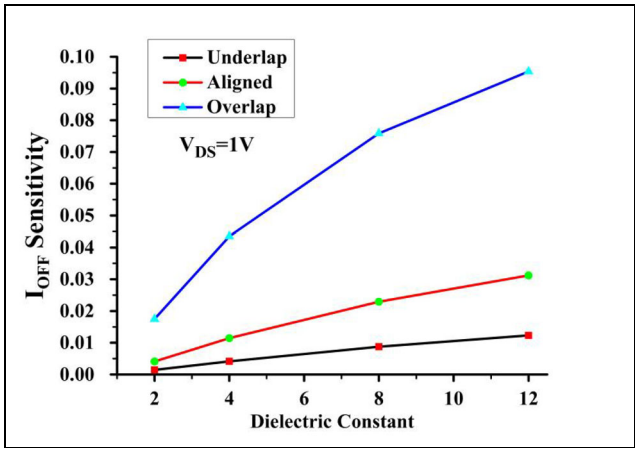


Fig. 12 I_{OFF} Sensitivity of DE-SBMOS based biosensor for UBS, OBS & ABS mode (V_{DS} = 1 V, k = 2, 4, 8, & 12)

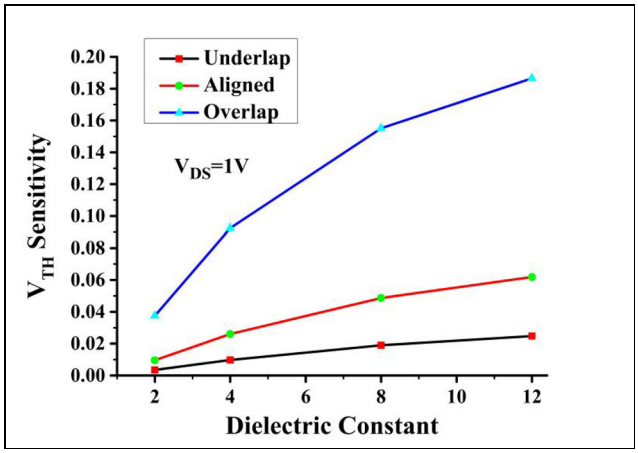


Fig. 13 V_{TH} Sensitivity of DE-SBMOS based biosensor for UBS, OBS & ABS mode (V_{DS} = 1 V, k = 2, 4, 8, & 12)

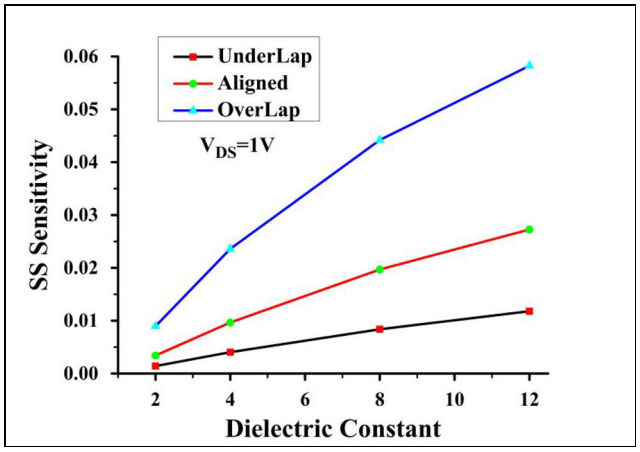


Fig. 14 SS Sensitivity of DE-SBMOS based biosensor for UBS, OBS & ABS mode (V_{DS} = 1 V, k = 2, 4, 8, & 12)

5 Conclusion

In this manuscript, we have investigated the sensing performance of the proposed DE-SBMOS based biosensor in terms of different sensitivity parameters for UBS, ABS, & OBS modes. The proposed device used both dielectric engineering and modulation with schottky barrier tunneling for the successful identification of different biomolecules. The simulation results show that out of three modes considered in the proposed work, the device with OBS mode turns out to be the optimized one. The OBS mode shows the higher sensitivity in terms of I_{ON} current and I_{DS} Sensitivity. The OBS mode shows 55.87% and 37.5% higher I_{ON} sensitivity than the UBS mode and ABS mode respectively, for k = 12. The OBS mode shows less sensitivity towards variation in I_{OFF} current, threshold voltage, sub threshold swing. So, OBS mode is less sensitive to leakage current and has a high drive current capability which helps in successful detection of the biomolecules. The proposed device exhibits enhanced sensitivity for detection of biomolecules. In addition, proposed device can be a strong candidate for future biosensor applications.

Availability of Data and Material All data included in this paper are available upon request by contact with the contact corresponding author.

Code Availability Not Applicable.

Authors' Contributions Rahul Singh[1], Shweta Kaim[2], Rani MedhaShree[3], Anil Kumar[4]: Writing, Original draft preparation, Formal Analysis, Investigation, Simulation, Data Curation.
Sumit Kale: Conceptualization, Methodology, Supervision.

Declarations

Conflict of Interest/Competing Interests The author has no conflicts of interest/competing interests to declare that are relevant to the content of this article.

Table 2 Sensitivity parameters for UBS, ABS, & OBS modes

Sensitivity Parameters	k-Value VValue	UBS Mode	ABS Mode	OBS Mode
V_{TH} Sensitivity	2	1.00	1.01	1.04
	4	1.01	1.03	1.09
	8	1.02	1.05	1.16
	12	1.02	1.06	1.19
I_{ON} Sensitivity	2	1.13	1.17	1.24
	4	1.46	1.60	1.88
	8	2.32	2.90	4.03
	12	3.46	4.90	7.84
I_{OFF} Sensitivity	2	1.00	1.00	0.98
	4	1.00	0.99	0.96
	8	0.99	0.98	0.92
	12	0.99	0.97	0.90
SS Sensitivity	2	1.00	1.00	0.99
	4	1.00	0.99	0.98
	8	0.99	0.98	0.96
	12	0.99	0.97	0.94
I_{DS} Sensitivity	2	0.13	0.17	0.24
	4	0.46	0.60	0.88
	8	1.32	1.90	3.03
	12	2.46	3.90	6.84

- The contents of this manuscript are not now under consideration for publication elsewhere;
- The contents of this manuscript have not been copyrighted or published previously.
- The contents of this manuscript will not be copyrighted, submitted, or published elsewhere, while acceptance by the Journal is under consideration.

Ethics Approval The authors declare that they have no known competing interest or personal relationship that could have appeared to influence the work reported in this paper.

Consent to Participate All the authors voluntarily agree to participate in this research work.

Consent to Publication All the authors provide their consent to publish the present work in the Silicon journal.

Human and Animal Rights This article does not contain any studies involving animals or human participants performed by any of the authors.

Financial Interests No financial interests.

References

1. Luong JHT, Male KB, Glennon JD (Sep. 2008) Biosensor technology: technology push versus market pull. Biotechnol Adv 26(5): 492–500. https://doi.org/10.1016/j.biotechadv.2008.05.007
2. Ghosh S, Chattopadhyay A, Tewari S (2020) Optimization of hetero-gate-dielectric tunnel FET for label-free detection and identification of biomolecules. IEEE transactions on Electron Devices, 1–8 67:2157–2164. https://doi.org/10.1109/ted.2020.2978499
3. Kanungo S, Chattopadhyay S, Gupta PS, Rahaman H (Mar. 2015) Comparative performance analysis of the dielectrically modulated full-gate and short-gate tunnel FET-based biosensors. IEEE Trans Electron Devices 62(3):994–1001. https://doi.org/10.1109/TED.2015.2390774
4. Kim S, Baek D, Kim JY, Choi SJ, Seol ML, Choi YK (2012) "A transistor-based biosensor for the extraction of physical properties from biomolecules," Appl. Phys. Lett., vol. 101, no. 7, Aug. Art. no. 073703, https://doi.org/10.1063/1.4745769
5. Sarkar D, Banerjee K (2012) "Proposal for tunnel-field-effect-transistor as ultra-sensitive and label-free biosensors," Appl. Phys. Lett., vol. 100,no. 14, Apr. Art. no. 143108, https://doi.org/10.1063/1.3698093
6. Kalra S, Kumar MJ, Dhawan A (Nov. 2016) Dielectric-modulated field effect transistors for DNA detection: impact of DNA orientation. IEEE Electron Device Lett 37(11):1485–1488. https://doi.org/10.1109/LED.2016.2613110
7. Narang R, Reddy KVS, Saxena M, Gupta RS, Gupta M (Oct. 2012) A dielectric-modulated tunnel-FET- based biosensor for label-free detection: analytical modeling study and sensitivity analysis. IEEE Trans. Electron Devices 59(10):2809–2817. https://doi.org/10.1109/TED.2012.2208115
8. Narang R, Saxena M, Gupta M (May 2015) Comparative analysis of dielectric-modulated FET and TFET-based biosensor. IEEE Trans Nanotechnol 14(3):427–435. https://doi.org/10.1109/TNANO.2015.2396899
9. Kanungo S, Chattopadhyay S, Gupta PS, Sinha K, Rahaman H (Jun. 2016) Study and analysis of the effects of SiGe source and pocket-doped channel on sensing performance of dielectrically modulated tunnel FET- based biosensors. IEEE Trans. Electron Devices 63(6):2589–2596. https://doi.org/10.1109/TED.2016.2556081
10. Larson JM, Snyder JP (May 2006) Overview and status of metal S/D Schottky-barrier MOSFET technology. in *IEEE Transactions on Electron Devices* 53(5):1048–1058. https://doi.org/10.1109/TED.2006.871842
11. Kale S, Kondekar PN (Sep 2017) Design and investigation of dielectric engineered dopant segregated Schottky barrier MOSFET with NiSi source/drain. IEEE Trans Electron Devices 64(11): 4400–4407
12. Jhaveri R, Woo JCS (2006) "Schottky tunneling source MOSFET design for mixed mode and analog applications," in Proc. 36th Eur. Solid State Device Res. Conf., Montreux, Switzerland, pp. 295–298,https://doi.org/10.1109/ESSDER.2006.307696
13. Kale S (2020) Investigation of dual metal gate Schottky barrier MOSFET for suppression of ambipolar current. IETE J Res. https://doi.org/10.1080/03772063.2020.1823250
14. Kale S, Kondekar PN (2015) Suppression of ambipolar leakage current in Schottky barrier MOSFET using gate engineering. Electron Lett 51(19):1536–1538. https://doi.org/10.1049/el.2015.0283
15. Kale S, Chandu MS (2021) Dual metal gate dielectric engineered dopant segregated Schottky barrier MOSFET with reduction in Ambipolar current. Silicon. https://doi.org/10.1007/s12633-020-00921-4
16. Hema Latha NK, Kale S (2020) Dielectric modulated Schottky barrier TFET for the application as label-free biosensor. Silicon 12:2673–2679. https://doi.org/10.1007/s12633-019-00363-7
17. Verma M, Sharma D, Pandey S, Nigam K, Kondekar PN (Jan. 2017) Performance comparison of single and dual metal dielectrically modulated TFETs for the application of label free biosensor. Superlattice Microst 101:219–227. https://doi.org/10.1016/j.spmi.2016.11.045
18. Ikarashi N, Watanabe K, Masuzaki K, Nakagawa T (2006) "Thermal stability of a HfO2/SiO2 interface," Appl. Phys. Lett., vol. 88, no. 10, Art. no. 101912, https://doi.org/10.1063/1.2182023
19. Pitera WJ, Falta M, Gunsteren WF (Jun. 2001) Dielectric properties of proteins from simulation: the effects of solvent, ligands, pH, and temperature. Biophys J 80(6):2546–2555. https://doi.org/10.1016/S0006-3495(01)76226-1
20. Hill GD, Walbergh DB, Frommelt PC (Apr. 2016) Efficacy of reconstituted oral chloral hydrate from crystals for echocardiography sedation. J Amer Soc Echocardiography 29(4):337–340. https://doi.org/10.1016/j.echo.2015.11.014
21. Kim J-Y, Ahn JH, Choi SJ, Im M, Kim S, Duarte JP, Kim CH, Park TJ, Lee SY, Choi YK (Mar. 2012) An underlap channel-embedded field-effect transistor for biosensor application in watery and dry environment. IEEE Trans Nanotechnol 11(2):390–394. https://doi.org/10.1109/TNANO.2011.2175006
22. Eckhardt S, Brunetto PS, Gagnon J, Priebe M, Giese B, Fromm KM (Jul. 2013) Nanobio silver: its interactions with peptides and bacteria, and its uses in medicine. Chem Rev 113(7):4708–4754. https://doi.org/10.1021/cr300288v
23. Im H, Huang X-J, Gu B, Choi Y-K (Jul. 2007) A dielectric-modulated field-effect transistor for biosensing. Nature Nanotechnol 2(7):430–434. https://doi.org/10.1038/nnano.2007.180
24. Kannan N, Kumar MJ (Dec. 2013) Dielectric-modulated impact-ionization MOS transistor as a label-free biosensor. IEEE Electron Device Lett. 34(12):1575–1577. https://doi.org/10.1109/LED.2013.2283858
25. Kannan N, Kumar MJ (Aug. 2015) Charge-modulated underlap I-MOS transistor as a label-free biosensor: a simulation study. IEEE

Trans. Electron Devices 62(8):2645–2651. https://doi.org/10.1109/TED.2015.2446612
26. Singh D, Pandey S, Nigam K, Sharma D, Yadav DS, Kondekar P (Jan. 2017) A charge-plasma-based dielectric-modulated junctionless TFET for biosensor label-free detection. IEEE Trans. Electron Devices 64(1):271–278. https://doi.org/10.1109/TED.2016.2622403
27. Moon D-I, Han J-W, Meyyappan M (Nov. 2016) Comparative study of field effect transistor based biosensors. IEEE Trans Nanotechnol 15(6):956–961. https://doi.org/10.1109/TNANO.2016.2615855
28. (2003). The International Technology Roadmap for Semiconductors (ITRS). [Online]. Available: http://www.itrs2.net
29. ATLAS Device Simulation Software, Silvaco Int., Santa Clara, CA, USA, 2012. [Online]. Available: http://www.silvaco.com
30. Lu X, Yang H, Ge Z, Lu N, Or DT, Kao C, Chang M, "Selective etching of silicon nitride, "U.S. Patent 8, 252,696, issued Aug. 2012, patent no.: US 8.252,696 B2
31. Starzynski J, "Selective hafnium oxide etchant. "U.S. Patent Application 10/938,191, filed March 16, Mar. 2006, pub. no.: US 2006/0054595 A1
32. Shamiryan D, Paraschiv V, Demand M, "Plasma composition for selective high-k etch. "U.S. Patent 7, 598,184, issued Oct. 2009, patent no. : US 7,598,184 B2
33. Brask JK, Shah U, Doczy ML, Kavalieros J, Chau RS, Turkot RB Jr, and Metz MV., "Selective etch process for making a semiconductor device having a high-k gate dielectric." U.S. Patent 7,037,845, issued May, 2006, patent no.: US 7,037,845 B2
34. Zheng Z, Yu X, Xie M, Cheng R, Zhang R, Zhao Y (Jul.2016) Demonstration of ultra-thin buried oxide germanium-oninsulator MOSFETs by direct wafer bonding and polishing techniques. Applied Physics Letters 109(2):023503. https://doi.org/10.1063/1.4955486
35. Sze SM (1981) Physics of semiconductor devices, 2nd ed. Wiley
36. Kumar P, Bhowmick B (Sep 2017) 2-D analytical modelling for electrostatic potential and threshold voltage of a dual work function gate Schottky barrier MOSFET. J Comput Electron 16(3):658–665
37. Venkatesh P, Nigam K, Pandey S, Sharma D, Kondekar PN (Sept. 2017) A dielectrically modulated electrically doped tunnel FET for application of label free biosensor. Superlattice Microst 109:470–479
38. Kumar P, Bhowmick B (May 2018) Scaling of dopant segregation Schottky barrier using metal strip buried oxide MOSFET and its comparison with conventional Device. Silicon 10(3):811–820
39. Kumar P, Bhowmick B, "2D analytical model for surface potential based electric field and impact of wok function in DMG SB MOSFET," Superlattices and Microstructures, vol. 109, pp. 805–814, Sep. 2017, https://doi.org/10.1016/j.spmi.2017.06.001 (Impact Factor : 2.117)
40. Kumar P, Bhowmick B, "A Physics Based Threshold Voltage Model for Hetero - Dielectric Dual Material Gate Schottky Barrier MOSFET", International Journal of Numerical Modelling: Electronic Networks, Devices and Fields. (Impact Factor: 0.622) Volume31, Issue5 September/October 2017
41. Kumar P, Bhowmick B, "Suppression of Ambipolar Conduction and Investigation of RF Performance Characteristics of Gate Drain Underlap SiGe Schottky Barrier Field Effect Transistor", Volume: 13 , Issue: 5 , 5 2018Micro & Nano Letters, (Impact Factor: 1.7) https://doi.org/10.1049/mnl.2017.0895 Page(s): 626–630 ISSN: 1750–0443 2017, 2018
42. Kumar P, Bhowmick B, "Source-Drain Junction Engineering Schottky Barrier MOSFETs and their Mixed Mode Application," Silicon, vol. 12, no. 4, pp. 821–830, Apr. 2020, https://doi.org/10.1007/s12633-019-00170-0. Silicon , 2019
43. Kumar P, Bhowmick BB, "Comparative Analysis of Hetero Gate Dielectric Hetero Structure Tunnel FET and Schottky Barrier FET with n+ Pocket Doping for Suppression of Ambipolar Conduction and Improved RF/Linearity," J. Nanoelectron. Optoelectron., p. 11, 2018.JNO, oct 2018. (Impact Factor: 0.78)
44. Kumar P, Bhowmick B, Vinod A "Impact of Ferroelectric on the Electrical Characteristics of Silicon-Germanium based heterojunction Schottky Barrier MOSFET", doi.org/10.1016/j.aeue.2019.05.030 ISSN-1434-8411 Volume 107, Pages 257–263 AEU, May 2019 (Impact Factor: 2.65)
45. Kumar P, Vinod A, Krishna D, Brinda B "analysis and Simulation of Schottky tunneling using Schottky barrier FET with 2-D analytical modeling", volume 107, Silicon, JAN 2021

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

2021 3rd International Conference on Signal Processing and Communication (ICPSC) | 978-1-6654-2864-4/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICSPC51351.2021.9451677

Embodying Current division transconductance reduction in M-CDTA and its application as low frequency filter

Mohit Bansal
Department of Electronics and Communication
Delhi Technological University
New Delhi, India
mohitbansal_2k17ec103@dtu.ac.in

Pranav Semwal
Department of Electronics and Communication
Delhi Technological University
New Delhi, India
pranavsemwal_2k17ec128@dtu.ac.in

Navneet Sagar
Department of Electronics and Communication
Delhi Technological University
New Delhi, India
navneetsagar_2k17ec108@dtu.ac.in

Neeta Pandey
Department of Electronics and Communication
Delhi Technological University
New Delhi, India
neetapandey@dce.ac.in

***Abstract*—This research paper brings forward an enhancement to existing Modified-Current Differencing Trans-Conductance Amplifier (M-CDTA) which works on subthreshold region in order to provide low g_m and make it more suitable to low frequency filter application. Here, current division g_m reduction technique is embodied in M-CDTA to provide even lower g_m. The proposed block is termed as New Modified-Current Differencing Trans-Conductance Amplifier (NM-CDTA) in context of this paper. The functional verification of the proposed block is done using Net-listing in ORCAD PSpice A/D simulation tool. A multiple input single output filter is also included in paper as an application of NM-CDTA and its operation is verified through simulations.**

Keywords— CDTA, Low-Transconductance, Subthreshold Region, Current Division.

I. INTRODUCTION

In recent years, analog building blocks (ABB) have acquired significant consideration because of their massive usage in different VLSI applications. Current mode ABB has acquired further acknowledgment in simple sign handling because of vast data transfer capacity, diminished force utilization, and expanded unique reach. The CDTA [1] is a versatile current processing ABB that may be used in variety of applications oscillators [2], memristor emulator realizations [3], capacitance multipliers [4], filters [5], inductor realizations [6], square wave generators [7] and Schmitt trigger [8]. The transconductance (g_m) value of CDTA is primarily used for controlling performance parameters of the application. A low g_m value is useful for biomedical signal processing i.e. filters having low pole frequencies [9], ramp signal generation for testing of analog-to-digital converters [10], neural networks [11], hearing aids, wearable gadgets and implantable pacemakers [12].

The organization of paper is in the following manner: Section 2 describes low-transconductance NM-CDTA block running in sub-threshold region. Section 3 presents the implementation and simulation results of the CM universal made to show application of proposed block. Section 4 presents conclusion.

II. PROPOSED NM-CDTA BLOCK

CDTA is a variable current processing Active Building Blocks (ABB) containing a Current Differencing Unit (CDU) cascaded with a transconductance amplifier (TA); it has five terminals. and the Systematic Operation of the CDTA block can be expressed using the port relationships provided in (1).

$$\begin{bmatrix} V_P \\ V_N \\ I_Z \\ I_{X+} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & g_m & 0 \end{bmatrix} \begin{bmatrix} I_P \\ I_N \\ V_Z \\ V_{X+} \end{bmatrix} \quad (1)$$

The functionality of CDTA can be described as follows:
The difference between input currents applied at the terminals P and N is passed through the Z terminal. The X+ terminal's output current is product of g_m (TA block) and voltage at the Z terminal.
The CDTA operating in saturation region provides g_m values in micro Siemens (μS). The pole frequencies in most filters are dependent on the g_m values of the ABB so a low g_m is required for low frequency filters. A modified CDTA is presented in [13] provides low g_m by redesigning [14] to operate in subthreshold region. This version of CDTA requires bias generation circuits for obtaining varying bias currents for electronic tuning of g_m of TA block This section describes the M-CDTA first and then NM-CDTA wherein current division method [15] is used for providing electronic tuning in place of varying bias current of TA block. Thus, the proposed NM-CDTA saves on bias current generation circuits.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:32:27 UTC from IEEE Xplore. Restrictions apply.

CMOS based implementation of MCDTA block is depicted in Fig. 1. The transistors M1-M14 and M24-M25 form the current differencing unit while M15-M23 and M26-M28 form TA block. It has three X+ terminals which carry the same current $g_m V_Z$.

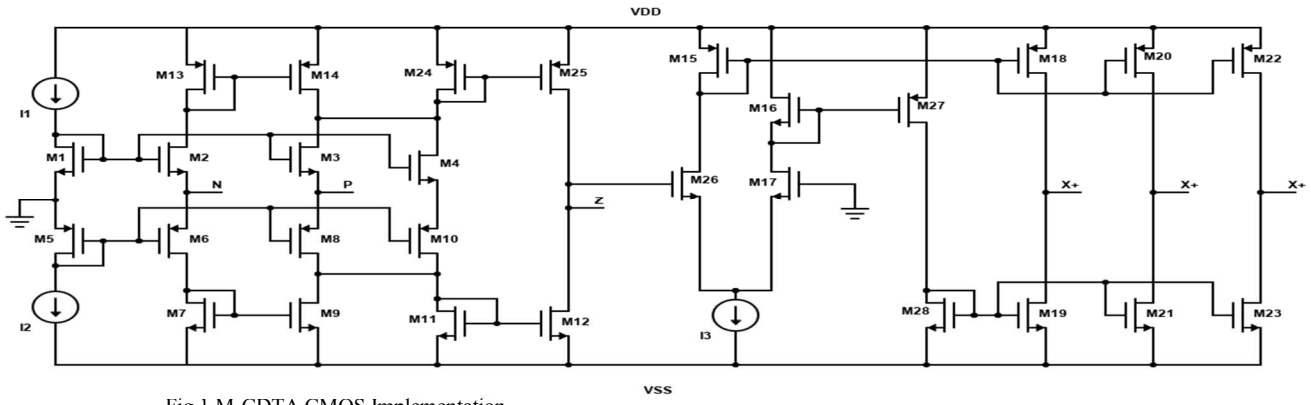


Fig.1 M-CDTA CMOS Implementation

The g_m of M-CDTA block is given by.

$$g_m = \frac{I_3}{2\eta V_t} \qquad (2)$$

where V_t corresponds to thermal voltage and η represents ideality factor.

The proposed NM-CDTA blocks' CMOS realization is represented in Fig. 2. It is a redesigned version of M-CDTA block [13] such that additional terminal can be embodied. A switch controlled current division technique [15] is applied to further reduce g_m which may be used to reduce pole frequency thus playing a vital role in biomedical areas. Through current division scheme, the PMOS load transistors M15 and M16 of differential amplifier in the TA block are replaced by N parallel symmetrical PMOS transistors which are controlled by N switches namely, $S15_i$, $S16_i$ (where i=1...N) which cut-or-make connection with supply voltage as per demand of current division blocks to be used.

So, now what happens is that, the current initially flowing through transistors (M15, M16), now flows through N parallel transistors placed on both the sides of differential amplifier

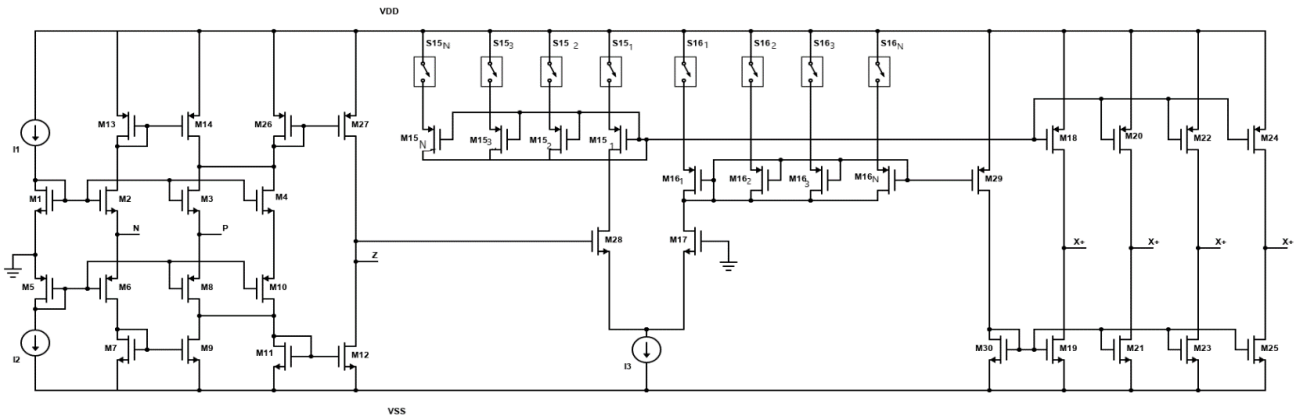


Fig.2 CMOS realization of NM-CDTA

depending on the condition of respective switches, which lead to reduction in load impedance. With this approach the output current can be expressed as (2) and hence g_m reduction (3). Thus, G_m reduces with increasing N.

$$I_{X+} = G_m V_Z \qquad (3)$$

$$G_m = \frac{gm}{N} \qquad (4)$$

An additional terminal X+ is introduced to provide flexibility in the design which is seen in symbolic representation of NM-CDTA in Fig.3.

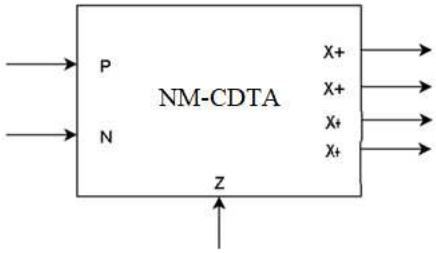


Fig 3: NM-CDTA Symbolic Representation

The verification of the proposed block is done on ORCAD PSpice A/D simulation tool. The aspect ratios of the transistors in Fig. 2 are enlisted in Table I, power supplies (VDD and VSS) are chosen to be 0.3V and -0.3V and the bias currents as I1 = I2 = I3 = 5nA. All the results are summarized in Table II.

To verify characteristic properties of NM-CDTA, simulations are carried out and the findings are placed in Fig. 4 and the observations are placed in Table II.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:32:27 UTC from IEEE Xplore. Restrictions apply.

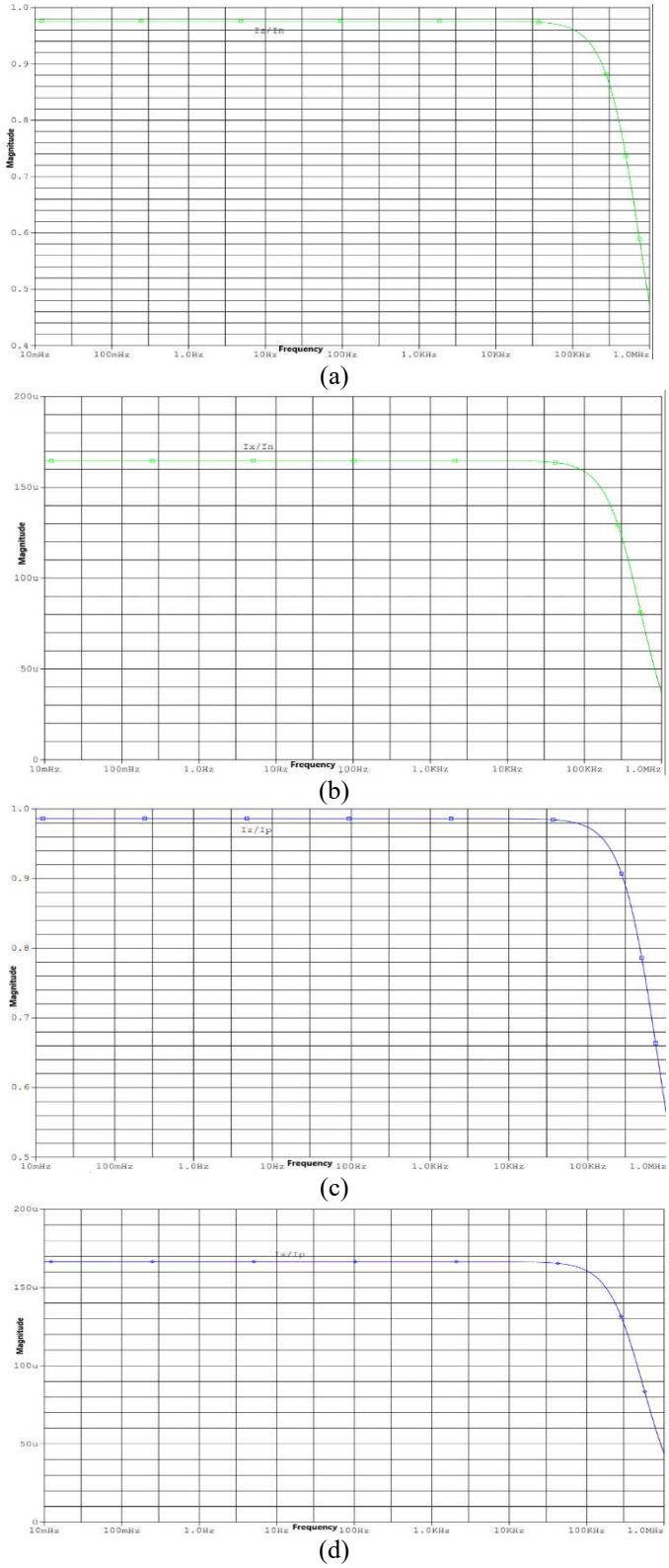


Fig. 4. Current transfer characteristics from (a) n to Z (b) n to X+ (c) p to Z (d) p to X+

TABLE I. Aspect Ratio of transistors

Transistors (Mi)	**Aspect Ratio**
1-4	600nm/500nm
5-10	2um/500nm
11-12	2um/1.5um
13-30, 15i, 16i (i=1, N)	700n/500n

TABLE II. Simulation Results

Conditions	**Values**
Supply Voltage	±0.3V
Technology Node	180nm
Bias Currents	5nA,5nA,5nA
I_z / I_p (mag, 3 dB freq)	0.986, 531.648KHz
I_z / I_n (mag, 3 dB freq)	0.976, 561.8KHz
I_{x+} / I_p (mag, 3 dB freq)	0.1665m, 349.148KHz
I_{x+} / I_n (mag, 3 dB freq)	0.1647m, 333.968KHz
Transconductance(g_m)	66.67 nS

Variation produced in transconductance of TA side of NM-CDTA with current division parameter N varying from 1-5 is presented in Fig. 5. It clearly shows that with increase in N there's decrease in Gm as depicted in (4) and results are summarized in Table III which are within the permissible limits.

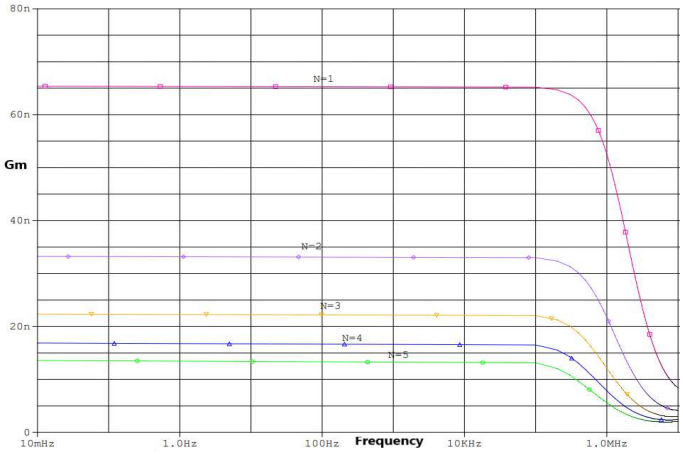


Fig.5 Variation in Gm with N

TABLE III. Summarized results of Fig.3

Value of N	Gm (nS)
1	65.64
2	33.4
3	22.71
4	17.2
5	14.47

III. LOW FREQUENCY CURRENT MODE FILTER

To verify the working of NM-CDTA block using values described in Table II, a CM universal filter [16] is represented as an application in Fig. 5. It uses two NM-CDTAs and two grounded capacitances. The output current (I_{out}) is expressed in terms of input currents (I_1, I_2 and I_3) as equation (5).

$$I_{out} = \frac{s\frac{Gm}{C_1}I_1 + \frac{Gm^2}{C_1C_2}I_2 + \left(s^2 + \frac{Gm}{C_1}s + \frac{Gm^2}{C_1C_2}\right)I_3}{s^2 + \frac{Gm}{C_1}s + \frac{Gm^2}{C_1C_2}} \quad (5)$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:32:27 UTC from IEEE Xplore. Restrictions apply.

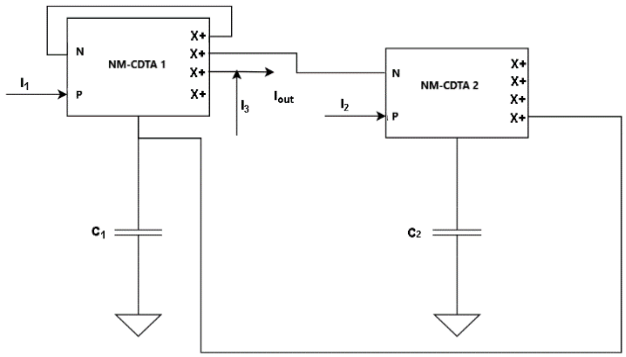


Fig. 5 CM Filter Implementation

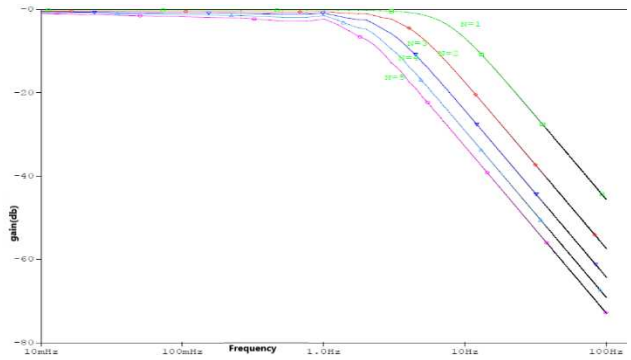


Fig. 6 LP Variation with N

The selection of input currents for achieving different filter responses is enlisted in Table IV along with their transfer function. The filter is characterized by pole frequency (6), bandwidth (7) and the quality factor (8).

$$\omega_o = \frac{G_m}{\sqrt{C_1 C_2}} \quad (6)$$

$$BW = \frac{C_1}{C_2} \quad (7)$$

$$Q_o = \sqrt{\frac{C_1}{C_2}} \quad (8)$$

It may be noted that pole frequency is reduced by decreasing G_m of the NM-CDTA block.

TABLE IV. Filter Responses

Filtered Response	Inputs			Transfer Function $D(S) = s^2 + \frac{G_m}{C_1}s + \frac{{G_m}^2}{C_1C_2}$
	I_1	I_2	I_3	
LP	0	I_{in}	0	$\frac{I_{out}}{I_{in}} = \frac{\frac{{G_m}^2}{C_1C_2}}{D(S)}$
HP	$-I_{in}$	$-I_{in}$	I_{in}	$\frac{I_{out}}{I_{in}} = \frac{s^2}{D(S)}$
BP	I_{in}	0	0	$\frac{I_{out}}{I_{in}} = \frac{\frac{G_m}{C_1}s}{D(S)}$
BR	$-I_{in}$	0	I_{in}	$\frac{I_{out}}{I_{in}} = \frac{s^2 + \frac{{G_m}^2}{C_1C_2}}{D(S)}$
AP	$-2I_{in}$	0	I_{in}	$\frac{I_{out}}{I_{in}} = \frac{s^2 - \frac{G_m}{C_1}s + \frac{{G_m}^2}{C_1C_2}}{D(S)}$

The values of C_1 and C_2 are taken as 1nF, 2nF respective to get Q_o=0.707. The effect of N is examined through the simulated magnitude response of different filters by varying values of N from 1 to 5. Since N is inversely proportional to Gm (4), hence pole frequency is decreased with increasing N which is clearly visible in Fig. 6 where variation in LP with N is shown. Table V enlists the pole frequencies for different values of N where practical and theoretical frequencies are close enough.

Table V. Results of pole frequency variation with N

Value of N	Low Pass Practical Pole Frequency (Hz)	Low Pass Theoretical Pole Frequency (Hz)
1	7.104	7.504
2	3.772	3.77
3	2.51	2.51
4	1.886	1.885
5	1.51	1.508

Figures 7, 8 and 9 represent variation in HP, Band Pass and Band Reject responses with N varying from 1-5. Fig. 10 present magnitude and phase plot of AP filter for N=4.

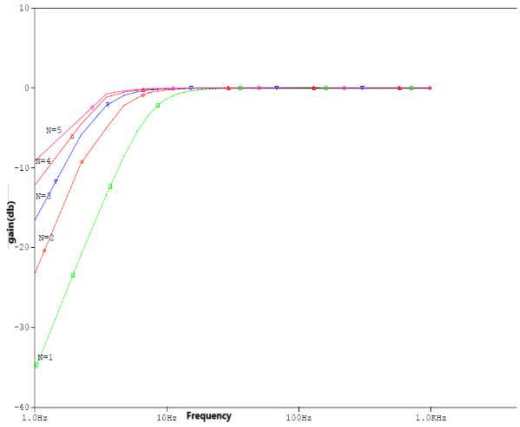


Fig.7 HP with variation in N

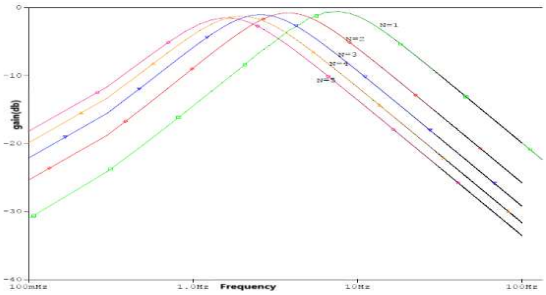


Fig.8 BP with variation in N

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:32:27 UTC from IEEE Xplore. Restrictions apply.

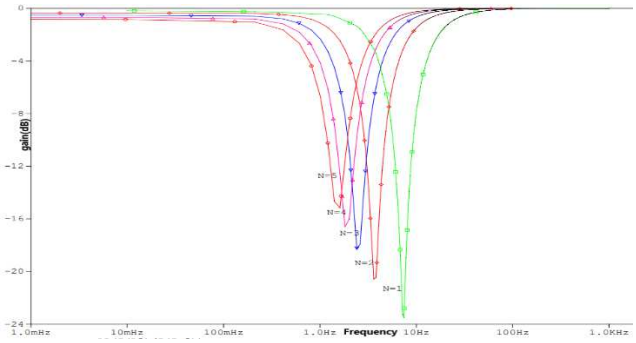


Fig.9 BR with variation in N

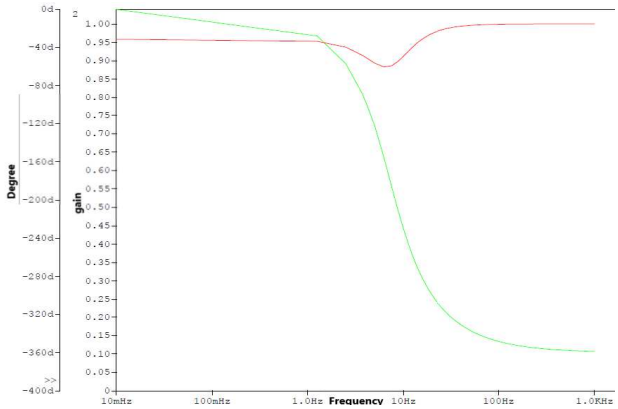


Fig.10 AP response

IV CONCLUSION

This research involves the realization of a low pass filter using CDTA as base block. Here we focus on a low transconductance New Modified-CDTA block (NM-CDTA), a version of M-CDTA with an additional output pin and controlled current divisional scheme which led to reduction in pole frequency. It is made to operate in the subthreshold region which results in a low value of transconductance, i.e. 17.2nS for current division parameter N=4. A low pass pole frequency 1.886Hz is obtained when the NM-CDTA block is utilized as a Universal CM filter. Further variation in different type of filter responses with N are also presented. All the simulation results are within permissible limits. Due to the grounding of all passive components of the filter, it is easy to fabricate along with appropriate input and output impedances.

REFERENCES

[1] D. Biolek, "CDTA-building block for current-mode analog signal processing", In Proceedings of the ECCTD, vol. 3, pp. 397-400, 2003.

[2] R. Das, B. Bhowmick, P. Paul, S. Karmakar, and K. Banerjee, "Current Differencing Transconductance Amplifier (CDTA) Based Current Mode Quadrature Oscillator", In Advances in Computer, Communication and Control, pp. 35-47, 2019.

[3] S. Gupta and S.K. Rai, "New Grounded and Floating Decremental/Incremental Memristor Emulators Based on CDTA and Its Application", Wireless Personal Communications, pp. 1-26, 2020.

[4] D. Biolek, J. Vavra, and A.Ü. Keskin, "CDTA-based capacitance multipliers", Circuits, Systems, and Signal Processing, vol. 38, no. 4, pp. 1466-1481, 2019.

[5] M. Kumngern and U. Torteanchai, "Four-input four-output currentmode multifunction filter using CDTAs", In 2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE), pp. 436-439, 2017.

[6] D. Prasad, Z. Haseeb, P. Mainuddin and M.W. Akram, "Realization of resistorless floating inductor using modified CDTA", Indian Journal of Pure & Applied Physics (IJPAP), vol. 57, no. 1, pp. 29-32, 2019.

[7] H.C. hien, "Current-mode square wave generators based on a single CDTA", International Journal of Electronics, vol. 104, no. 9, pp. 1589-1605, 2017.

[8] K. Nagalakshmi, A. Srinivasulu, C. Ravariu, V. Vijay and V.V.S.V. Krishna, "A Novel Simple Schmitt Trigger Circuit using CDTA and its application as a Square-Triangular Waveform Generator", Journal of Modern Technology and Engineering, vol. 3, no. 3, pp. 205-216, 2018.

[9] A. A. Fedotov, "Selection of parameters of bandpass filtering of the ECG signal for heart rhythm monitoring systems.", Biomedical Engineering, vol. 50, no. 2, pp. 114-118, 2016.

[10] P. Kinget, M. Steyaert, and J. Van der Spiegel, "Full analog CMOS integration of very large time constants for synaptic transfer in neural networks", Analog Integrated Circuits and Signal Processing, vol. 2, no. 4, pp. 281-295, 1992.

[11] M.R. DeWitt, G.F. Gross Jr and R. Ramachandran, "Built-in self-test for analog to digital converters", U.S. Patent 5,132,685, 1992.

[12] S. S. Rajput and S.S. Jamuar, "Low voltage analog circuit design stechniques", IEEE Circuits and System Magazine, vol. 2, no. 1, pp. 24-42, 2002.

[13] V. Bhanoo, A. Gangal and N. Pandey, "Subthreshold Low-Transconductance M-CDTA based Low Frequency Current Mode Universal Filter", International Journal of Engineering Trends and Technology, vol. 68. pp.16-22, 2020.

[14] M. Kumngern, K. Khwama and S. Junnapiya, "Three-input single-output current-mode universal filter using a MCDTA," In 2013 Eleventh International Conference on ICT and Knowledge Engineering, pp. 1-4, 2013.

[15] A. Veeravalli, E. Sánchez-Sinencio and J. Silva-Martínez, "Transconductance amplifier structures with very small transconductances: A comparative design approach.", IEEE Journal of Solid-State Circuits, vol. 37, no. 6, pp. 770-775, 2002.

[16] J. Jin and C. Wang, "Current-mode universal filter and quadrature oscillator using CDTAs", Turkish Journal of Electrical Engineering & Computer Sciences, vol. 22, no. 2, pp. 276-286, 2014.

[17] A.Ü. Keskin, D. Biolek, E. Hancioglu and V. Biolková, "Currentmode KHN filter employing current differencing transconductance amplifiers", AEU-International Journal of Electronics and Communication, vol.60, no. 6, pp. 443-446,2006.

[18] A. Gangal, V. Bhanoo and N. Pandey, "Low Transconductance M-CDTA Block and its Application in Biomedical Signal Processing", In International Conference on Electronics and Sustainable Communication Systems (ICESC), pp. 1101-1107, 2020.

[19] Y. Lakys and A. Fabre , ""Shadow filters – new family of second-order filters", Electronics Letters, vol. 46, no.4, 2010.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:32:27 UTC from IEEE Xplore. Restrictions apply.

Biointerface Research in Applied Chemistry
Platinum Open Access Journal (ISSN: 2069-5837)

Article
Volume 12, Issue 2, 2022, 1670 - 1681
https://doi.org/10.33263/BRIAC122.16701681

Evaluation of Different Phytochemicals Against BRCA2 Receptor

Karanpreet Singh Bhatia [1,†], Saksham Garg [2,†], Ashutosh Anand [2,†], Arpita Roy [3,*]

[1] National University of Singapore, Singapore
[2] Delhi Technological University, India
[3] Department of Biotechnology, School of Engineering & Technology, Sharda University, Greater Noida, India
[†] the authors contributed equally;
[*] Correspondence: arbt2014@gmail.com (A.R.);

Scopus Author ID 57203962973

Received: 15.04.2021; Revised: 20.05.2021; Accepted: 23.05.2021; Published: 9.06.2021

Abstract: Cancer is one of the most common diseases and remains a leading cause of death worldwide. Breast cancer is one of the major causes of mortality in women. It is one of the most frequent cancers among others. Due to its high mortality, it requires effective prevention. Radiotherapy, immunotherapy, and chemotherapy are the most common methods used to treat breast cancer, but these techniques adversely affect healthy cells. Therefore, an alternative therapy is required to overcome this problem. Phytochemicals from different plant sources are an alternative way for curing infectious diseases and now they are being explored for their anti-cancer therapeutics. In the present study, the BRCA2 receptor protein, which is actively involved in breast cancer expression, was chosen for the molecular docking analysis. This study focuses on evaluating phytochemical compounds from 5 different plants against BRCA2 breast cancer tumor receptors. Prior to computational analysis, a theoretical ADME study was used to rule out a few compounds. Virtual screening of these compounds identified five-hit molecules, which could be further explored in the drug discovery pipeline. Molecular docking analysis revealed that isocolumbin possesses maximum negative binding energy against BRCA2 receptors. These hits were found to be well in range in the bioavailability radar test as well.

Keywords: Phytochemicals; breast cancer; *in-silico* screening; docking; isocolumbinm; bioavailability.

© 2021 by the authors. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Breast cancer, one of the most frequent cancers among women, affected an estimated 2.1 million women in 2018, which translates to one new case every 18 seconds [1]. Approximately 15% of women died from breast cancer [2-3]. Regardless of the income level, the burden of breast cancer is increasing in women due to the increasing and aging population [3-4]. The death rate is also varied among the cancer subtypes, with HER2 positive having the highest death rate followed by TNBC, Luminal A and Luminal B subtypes [5]. In high-income countries, cancer is often diagnosed at early stages, whereas it is diagnosed at later stages in lower- and middle-income countries. Therefore, the prognosis in high-income countries is better than that in low- and middle-income countries [4]. Some studies also reported that breast cancer is present earlier age (generally 40-50 years of age) in Asian women as compared to Western region countries' women (generally between 60-70years of age) [7-9].

The genetic mutation in the breast cancer susceptibility gene BRCA2 confers a predisposition to individuals' early onset of breast cancer. BRCA2 encodes BRCA2 protein, whose function is to repair DNA double-strand breaks via homologous recombination [10,11]. One of the crucial steps in homologous recombination is recruiting RAD51 recombinase to the exposed single-stranded DNA overhangs. RAD51 forms highly ordered nucleoprotein filaments that mediate the search for homologous sequences on the undamaged sister chromatids, followed by DNA strand exchange [12]. The defective cells for BRCA2 are sensitive to DNA interstrand cross-linking agents and show a decreased efficiency of DNA double-stranded break repair. They also have the inability to promote efficient restart of stalled replication forks [13,14].

The development of graphics processing units and bioinformatics makes it an innovative tool to design new drugs [15, 16]. It reduces the monetary and time costs that generally go into experiments and trials. Docking combines the chemical and physical principles with scientific calculation algorithms to generate characteristic interactions between receptor and ligand molecule [17]. Natural compounds are one of the potential sources of bioactive compounds. Plant-derived compounds possess a wide range of therapeutic activities, including anti-cancer, antibacterial, antidiabetic, etc. [18-22]. Anti-cancer activity of various plant-derived compounds has been reported in various studies [23-26]. Therefore, this study focuses on evaluating phytochemical compounds from 5 different plants, namely - Tinospora cordifolia, Ocimum tenuiflorum, Podophyllum hexandrum *Andrographis paniculate,* and *Beta vulgaris,* against BRCA2 breast cancer tumor receptor.

2. Materials and Methods

2.1. Protein/Macromolecule.

PALB2/BRCA2 complex was used in this study. The 3-D structure (PDB id: 3EU7) of the complex was retrieved from RCSB PDB in PDB file format (https://www.rcsb.org/). PALB2/BRCA2 complex has 6 amino acid residues in the active site, namely: TYR1064, PHE1071, LEU1092, TYR1108, LEU1142, and LEU1143. The active site of the complex was involved in binding with GOL (Glycerol) inhibitor, as seen in Figure 1 [27]

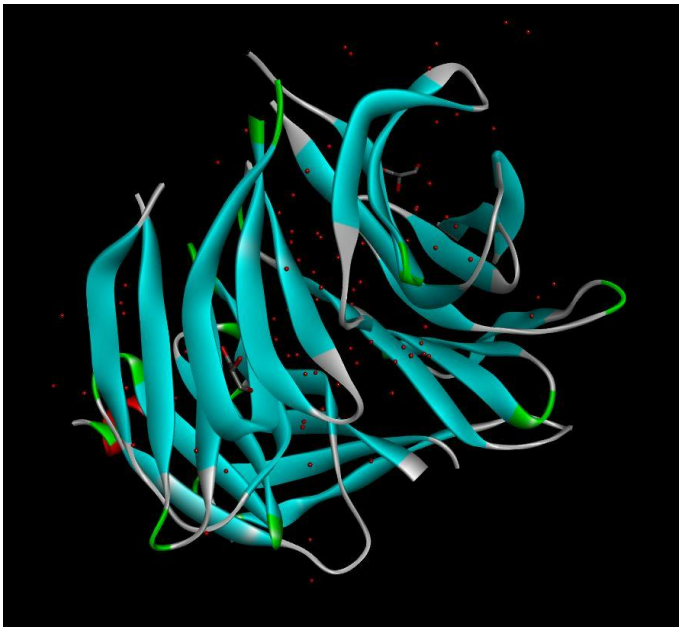


Figure 1. 3-D structure of PALB2/BRCA2 complex with GOL (Glycerol) inhibitor.

2.2. Ligands.

A total of 63 bioactive compounds were used as ligands. These selected compounds are from 5 different plants, namely *Tinospora cordifolia*, *Ocimum tenuiflorum*, *Podophyllum hexandrum*, *Andrographis paniculate*, and *Beta vulgaris*. Ligand structures were taken from PubChem databank in .sdf format. However, for docking purposes, these structures were converted to .pdb format using Biovia Discovery Studio Visualizer v20.1.0.19295 (https://pubchem.ncbi.nlm.nih.gov/).

2.3. ADME analysis.

A web tool named SwissADME was used for the initial screening of the shortlisted compound based on Lipinski's rule of five. To qualify as a potential ligand, this role states that a compound should have molar weight < 500Da; molar refractivity between 40 to 130 m3mol-1; high lipophilicity, i.e., logP < 5; hydrogen bond donor less than 5 and acceptors less than10. Any compound violating two or more of the parameters mentioned above was ruled out from further studies.

2.4. Molecular docking.

Autodock 4.2 was used to get the docked protein-ligand complex. The structure of every ligand and protein underwent optimization before docking. From the 3-D structure of the protein, water molecules, as well as inhibitors, were removed. Further, the addition of polar hydrogen bonds, Kollman charges and Gasteiger charges concluded the step of protein and ligand preparation. A grid box of dimension 54×40×40 with the spacing of 0.375 Å was made around the protein's binding site. The grid was centered at 6.751, -7.741 and 7.743 in x, y and z, respectively. The genetic algorithm was used as the search parameter and Lamarckian GA was used to handle the output. Docking Log File (DLG) was used for further analysis. The resulting DLG file reports a total of 10 conformations for each ligand. However, the most stable conformation, i.e., the conformation with the most negative binding energy, was selected as the final docking complex. This complex was converted to a 2-D structure to study the protein and ligand interactions at the binding site.

2.5. Bioavailability radar.

The Drug-likeliness of the compounds having binding energy less than the control was further analyzed. Six physiochemical properties: size, polarity, lipophilicity, solubility, flexibility, and saturation were considered to prepare a bioavailability radar via the SwissADME web tool. The shaded region shown in the radar chart represents the optimal values for the aforementioned physicochemical properties. Compounds exhibiting large deviations from these values suggest that the compound is not orally bioactive.

2.6. PASS web server.

PASS stands for Prediction of Activity Spectra for Substances. It is a tool used to evaluate the biological potential of an organic drug-like molecule. PASS reports predictions of various biological activities based on the structure of the compound. It calculates the probability P_a and P_i, which describes whether the given compound belongs to the active subclass or inactive subclass, respectively.

3. Results and Discussion

3.1. ADME analysis.

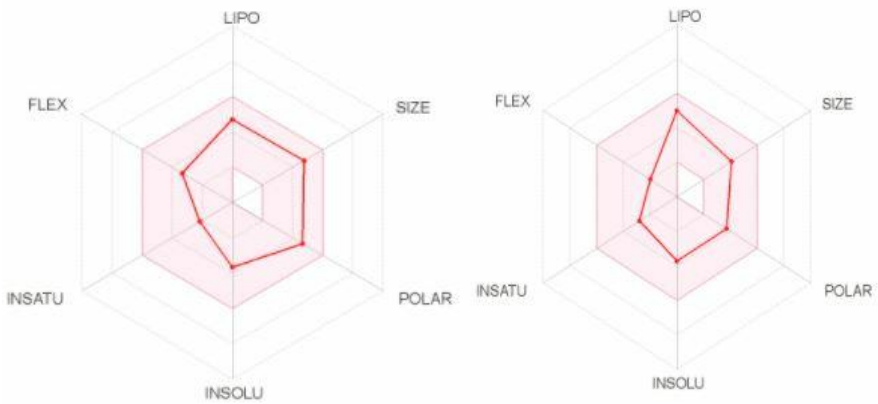
In order to estimate the drug-likeliness, Lipinski's rule was used on all the selected 63 bioactive compounds. Lipinski's rule is used to evaluate a bioactive compound's drug-likeliness and assist in ruling out few compounds based on their chemical and physical properties. However, Lipinski's rule only reports whether a compound is likely to be orally active in humans. Therefore, a compound qualifying the parameters doesn't make it a drug molecule. Compounds that violate two or more than two parameters were not considered for further studies. 18 bioactive compounds were withheld from further studies. Table 1 enlists the compounds considered for further studies.

Table 1. ADME analysis of bioactive compounds.

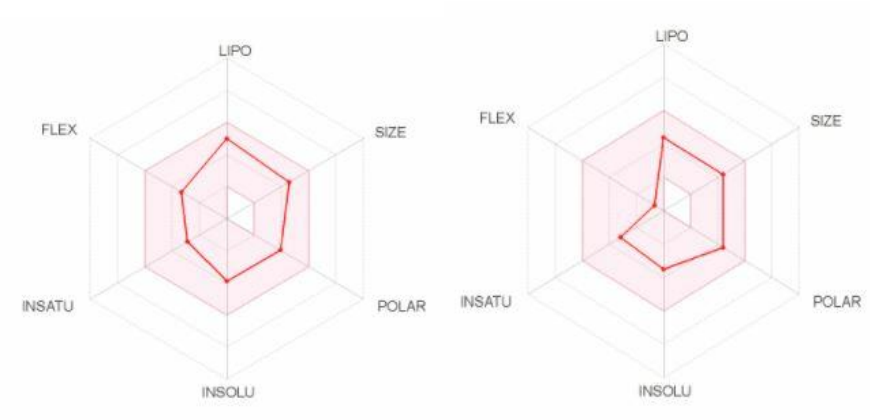
Compound Name	PubChem ID	Lipophilicity (LogP < 5)	H Bond Donor (<5)	H Bond Acceptor (<10)	Molar Refractivity (40-130)	Molecular Weight (<500 Da)	Violation(s)
Berberin	2353	3.6	0	4	94.87	336.4	0
Choline	305	3.6	1	1	29.69	104.17	1
Magnoflorine	73337	2.7	2	4	101.87	342.4	0
Tinosporin	122206355	1.1	2	8	97.66	406.4	0
Palmatine	19009	3.7	0	4	101.8	352.4	0
Isocolumbin	24721165	2.2	1	6	90.1	358.4	0
Aporphine	114911	3	0	1	78.97	235.32	0
Jatrorrhizine	72323	3.4	1	4	97.33	338.4	0
Tetrahydropalmatine	5417	3.2	0	5	103.99	355.4	0
Palmarin	442068	1.3	1	7	89.54	374.4	0
Tembetarine	167718	3	2	4	102.87	344.4	0
Eugenol	3314	2	1	2	49.06	164.2	0
Methyl Eugenol	7127	2.5	0	2	53.53	178.23	0
Cirsilineol	162464	2.9	2	7	91.44	344.3	0
Cirsimartin	1888323	2	2	6	84.95	314.29	0
Isothymusin	630253	2.6	3	7	86.97	330.29	0
Apigenin	5280443	1.7	3	5	73.99	270.24	0
Rosemeric Acid	5281792	2.36	5	8	91.4	360.31	0
Luteolin	5280445	2.53	4	6	76.01	286.24	0
Bornyl Acetate	6448	4.3	0	2	56.33	196.29	0
Myrtenal	61130	2.98	0	1	45.42	150.22	0
Neral	643779	3.03	0	1	49.44	152.23	0
Alpha Elemene	80048	5	0	0	70.42	204.35	0
Podophyllotixin	10607	2.01	1	8	103.85	414.41	0
4-dimethylpodophyllotoxin	122667	1.68	2	8	99.38	400.38	0
Quercetin	5280343	1.54	5	7	78.03	302.24	0
Kaempferol	5280863	1.9	4	6	76.01	286.24	0
Andrographolide	5318517	2.16	3	5	95.21	350.45	0
Isopropylideneandrographolide	71589898	3.17	1	5	107.41	390.51	0
14-deoxy-11,12-didehydroandrographolide	5708351	3.23	2	4	93.58	332.43	0
Dehydroandrographolide	6473762	3.38	2	4	93.58	332.43	0
Moslosooflavone	188316	3.32	1	5	82.93	298.29	0
19-O-acetyl-14-deoxy-11,12-	46179874	3.26	1	5	103.32	374.47	0

Compound Name	PubChem ID	Lipophilicity (LogP < 5)	H Bond Donor (<5)	H Bond Acceptor (<10)	Molar Refractivity (40-130)	Molecular Weight (<500 Da)	Violation(s)
didehydroandrographolide							
(8S,12R)-Isoandrographolide	101569019	2.34	2	5	93.55	350.45	0
14-acetylandrographolide	71589914	2.74	2	6	104.95	392.49	0
Azatoxin	125383	2.94	2	5	106.4	380.39	0
Betaine	247	0	0	2	28.35	117.15	1
Ferulic Acid	445858	1.51	2	4	51.63	194.18	0
Caffeic Acid	689043	1.15	3	4	47.16	180.16	0
p-Coumaric Acid	637542	1.46	2	3	45.13	164.16	0
Syringic Acid	10742	1.04	2	5	48.41	198.17	0
Rhamnetin	5281691	1.87	4	7	82.5	316.26	0
Rhamnocitirin	5320946	2.22	3	6	80.48	300.26	0
Betaxanthin	135926572	1.38	2	7	94.99	358.35	0
Orientin	5281675	0	8	11	108.63	448.38	2
Vicenin-2	3084407	-2.26	11	15	139.23	594.52	4
Ursolic Acid	64945	7.34	2	3	136.91	456.7	2
Podophyllotoxin glucoside	161177	0.42	4	13	136.23	576.55	2
Etoposide	36462	0.6	3	13	139.11	588.56	3
Teniposide	452548	1.24	3	13	156.66	656.65	3
NK611	9852558	2.12	2	13	157.42	652.09	3
GL331	148091	3.88	2	9	135.86	520.49	2
TOP-53	177859	3.02	1	9	137.66	512.59	2
Etoposide Phosphate	6918092	0	4	16	149.76	668.54	3
Tafluposide	9877073	4.84	2	30	220.84	1116.71	3
Beta-Sitosterol	222284	9.34	1	1	133.23	414.71	2
Stigmasterol	5280794	8.56	1	1	132.75	412.69	2
Rutin	5280805	0	10	16	141.38	610.52	4
Astragalin	5282102	0.72	7	11	108.13	448.38	2
Oleanolic Acid	10494	7.49	2	3	136.65	456.7	2
Beta-Carotene	5280489	13.54	0	0	184.43	536.87	3
Lutein	5281243	11.01	2	2	186.76	568.87	3

(A)14acetyl **(B)14deoxy**



(C)190Acetyl **(D)Isocolumbin**



(E)Isoprpoyl

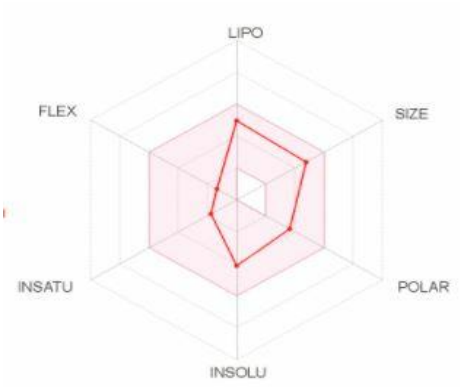


Figure 2. Bioavailability radar representation of five selected hits. The colored zone represents the preferable physicochemical space for oral bioavailability. LIPO stands for Lipophilicity: -0.7 <XLOGP3<+5.0. SIZE: 150 g/mol<TPSA< 130Å2. INSOLU stands for Insolubility: 0<LogS (ESOL) < fraction Csp3 < Number of rotatable bonds <9.

Table 2. Virtual screening results of the drug library. The top five molecules were further analyzed for their bioavailability using the pass webserver.

Compounds	Binding Energy	Ligand Efficiency	Inhibition constants (uM)	Intermolecular energy	vdw H bond desolvation energy
Isocolumbin	-8.73	-0.34	0.39684	-9.33	-9.37
19-O-acetyl-14-deoxy-11,12-didehydroandrographolide	-8.5	-0.31	0.5926	-10.29	-10.18
Isopropylideneandrographolide	-7.99	-0.29	1.39	-8.88	-8.79
14-acetylandrographolide	-7.61	-0.27	2.64	-9.7	-9.58
14-deoxy-11,12-didehydroandrographolide	-7.49	-0.31	3.25	-8.98	-8.92
Dehydroandrographolide	-7.35	-0.31	4.07	-8.84	-8.73
Berberin	-7.32	-0.29	4.3	-7.92	-7.86
Aporphine	-7.32	-0.41	4.32	-7.32	-6.84
Andrographolide	-7.32	-0.29	4.34	-9.11	-8.94

Compounds	Binding Energy	Ligand Efficiency	Inhibition constants (uM)	Intermolecular energy	vdw H bond desolvation energy
(8S,12R)-Isoandrographolide	-7.28	-0.29	4.64	-8.47	-8.37
Azatoxin	-7.28	-0.26	4.6	-8.47	-8.43
Palmatine	-7.02	-0.27	7.21	-8.21	-8.04
4-demethylpodophyllotoxin (CID: 122667)	-6.85	-0.24	9.46	-8.35	-8.31
Tetrahydropalmatine	-6.83	-0.26	9.84	-8.02	-7.63
Magnoflorine	-6.68	-0.27	12.8	-7.87	-7.42
Tinosporin	-6.66	-0.23	13.2	-8.15	-7.93
Palmarin	-6.54	-0.24	16.13	-7.13	-7.17
Podophyllotoxin (CID: 10607)	-6.53	-0.22	16.23	-8.03	-7.98
Apigenin (CID: 5280443)	-6.51	-0.33	17.03	-7.7	-7.53
Cirsilineol	-6.46	-0.26	18.37	-8.25	-7.89
Alpha Elemene (CID: 80048)	-6.42	-0.43	19.84	-7.01	-7
Cirsimaritin	-6.38	-0.28	21.03	-7.87	-7.8
Jatrorrhizine	-6.25	-0.25	26.06	-7.45	-7.39
Tembetarine	-6.24	-0.25	26.65	-8.03	-7.45
Moslosooflavone	-6.24	-0.28	26.75	-7.43	-7.42
Kaempferol (CID: 5280863)	-6.23	-0.3	27.28	-7.72	-7.57
Rhamnocitrin	-5.96	-0.27	42.57	-7.45	-7.38
Luteolin (CID: 5280445)	-5.95	-0.28	43.63	-7.44	-7.18
Quercetin (CID: 5280343)	-5.91	-0.27	46.71	-7.7	-7.43
Isothymusin	-5.74	-0.24	61.7	-7.53	-7.45
Bornyl acetate (CID: 6448)	-5.67	-0.41	69.99	-6.26	-6.25
Rhamnetin	-5.48	-0.24	96.05	-7.27	-7.07
Betaxanthin	-5.41	-0.21	108.89	-7.79	-8.43
Rosmarinic Acid (CID: 5281792)	-5.35	-0.21	119.52	-8.93	-8.87
Myrtenal (CID: 61130)	-5.34	-0.49	121.98	-5.64	-5.64
Methyl Eugenol	-4.42	-0.34	572.61	-5.62	-5.6
Neral (CID: 643779)	-4.34	-0.39	659.03	-5.53	-5.52
Eugenol	-4.27	-0.36	744.04	-5.46	-5.41
Caffeic Acid	-3.95	-0.3	1280	-5.44	-5.45
p-coumaric acid	-3.89	-0.32	1410	-5.08	-5.29
ferulic acid	-3.72	-0.27	1860	-5.22	-5.77
Syringic acid	-3.33	-0.24	3600	-4.83	-5.16
Choline	-2.78	-0.4	9210	-3.67	-3.3
Betaine	-2.56	-0.32	13230	-3.16	-2.59

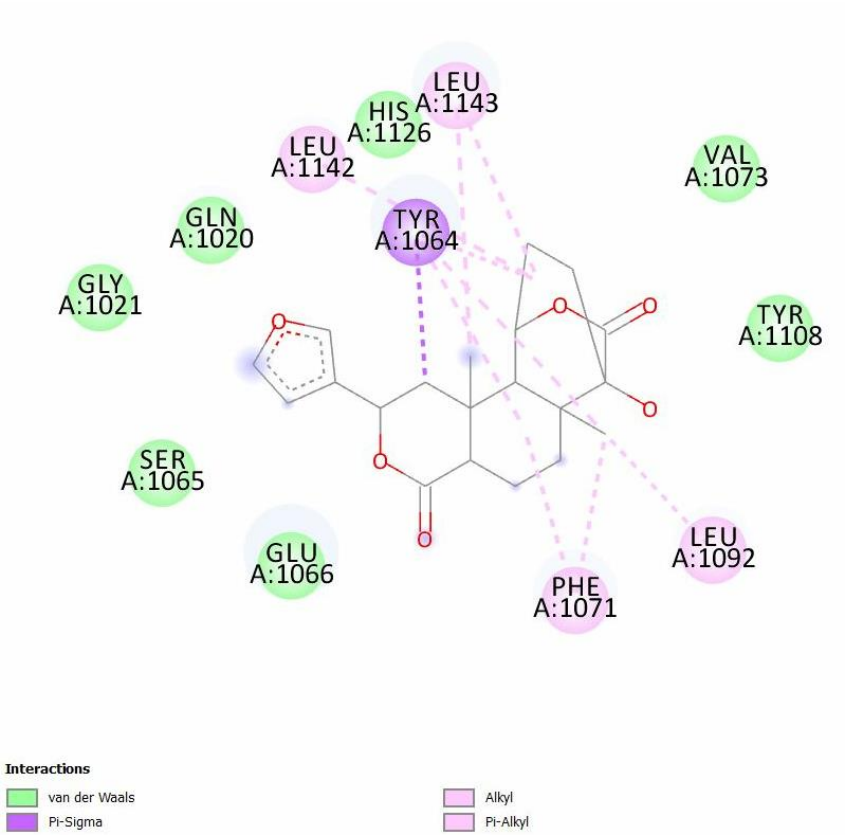


Figure 3. Interaction of Isocolumbin with residues of 3EU7.

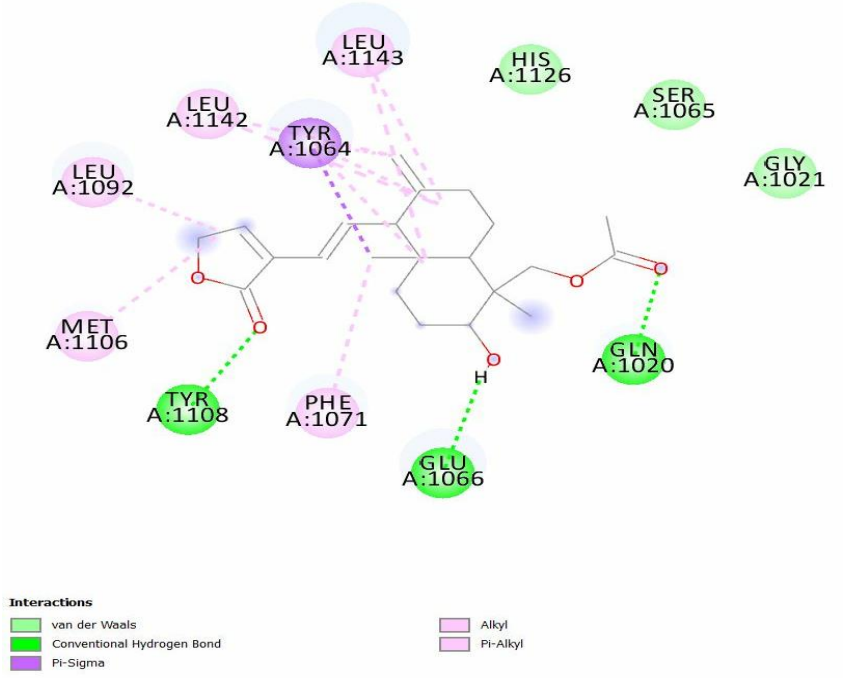


Figure 4. Interaction of 19-O-acetyl-14-deoxy-11,12-didehydroandrographolide with residues of 3EU7.

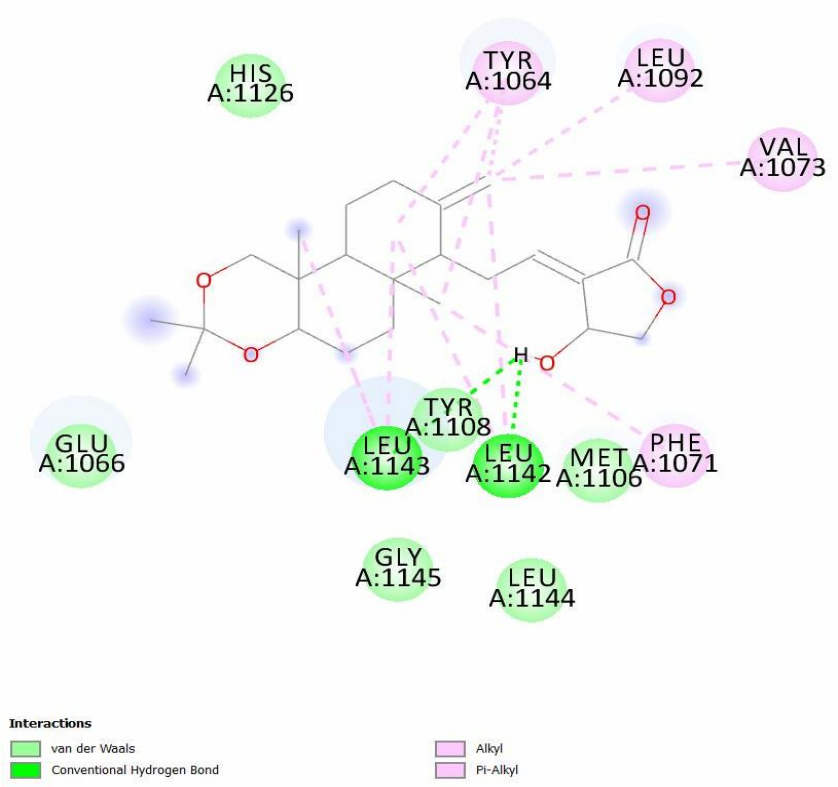


Figure 5. Interaction of Isopropylideneandrographolide with residues of 3EU7.

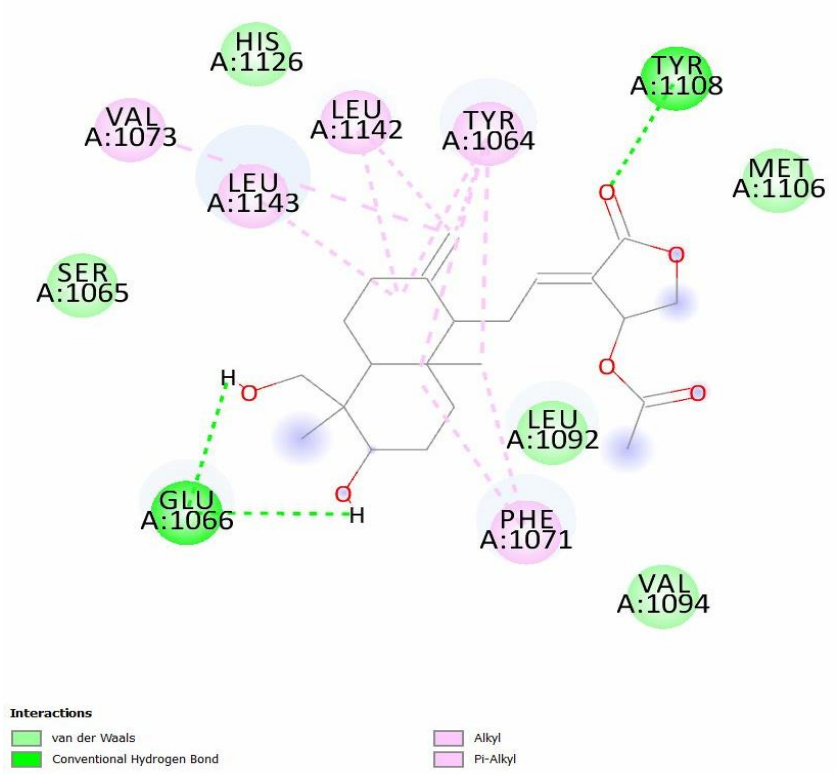


Figure 6. Interaction of 14-acetylandrographolide with residues of 3EU7.

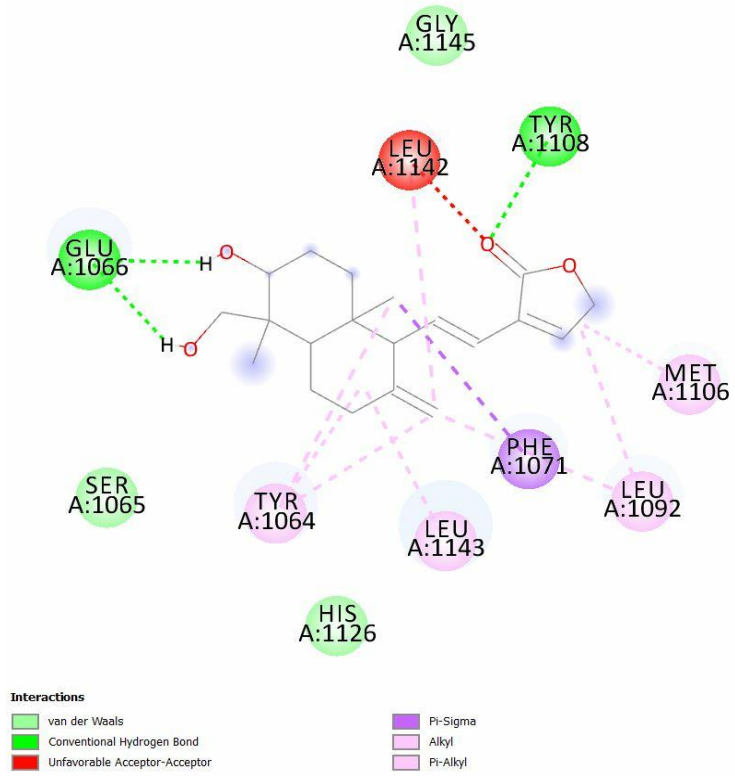


Figure 7. Interaction of 14-deoxy-11,12-didehydroandrographolide with residues of 3EU7.

Table 3. Results of Bioavailability prediction of the selected compounds.

Compound	Biological activity	Pa	Pi
Isocolumbin	Antineoplastic	0.882	0.005
19-O-acetyl-14-deoxy-11,12-didehydroandrographolide	Antineoplastic	0.904	0.005
Isopropylideneandrographolide	Antineoplastic	0.954	0.004
14-acetylandrographolide	Antineoplastic	0.959	0.004
14-deoxy-11,12-didehydroandrographolide	Antineoplastic	0.914	0.005

3.2. Pa- the probability that the compound is active and Pi- the probability that the compound is inactive.

The compounds which passed the Lipinski test were virtually screened using AutoDock and results are reported in Table 2. The top five hits with binding energy less than or equal to -7.49 kcal/mol were selected for bioavailability prediction, which results have been shown in Table 3. Drug likeness of these five hits was also evaluated using SWISSADME and the results are given in Figure 2. Figure 2 demonstrates that these five hits possess high drug-likeness as all of these compounds show ADME properties under the range of standard oral drugs. Results in Table 3 predict the biological activity of these hits. For a molecule to be biologically active, its Pa should be greater than Pi and all the five hits satisfied this condition. As reported by Baseer *et al*., [28], the amino acid residues which show interaction in BRCA2 are TYR1064, PHE1071, LEU1092, TYR1108, LEU1142, and LEU1143, with PHE1071 being the most important one. The top five hits identified in this study also show interaction with these residues, as depicted in Figure 3-7, emphasizing the relevance of these natural compounds. In all of these compounds, pi-alkyl interaction was majorly found in between the ligand molecule and PHE1071. Interaction with other residues is also quite similar in these hits, demonstrating

the importance of these interactions as all of these molecules have relatively better binding energy and constant inhibitory values.

4. Conclusions

Many researchers have explored BRCA2 and various drugs have been reported for the treatment of breast cancer manifestations with BRCA2 complexities. Side effects of the reported drugs, a resurgence of the tumor and resistance to the chemotherapy regime are some reasons for the ongoing research for the development of new compounds to tackle these issues. Natural products could answer these problems; therefore, in this study, the top five hits which we have identified after the Lipinski test, virtual screening and bioavailability prediction can further be analyzed in MD simulations and could serve as promising lead molecules which could be optimized and finally tested in clinical settings for their efficacy and safety.

Funding

Dr. Arpita Roy is thankful to Sharda University for providing seed fund (Seed fund-4 2001 (SUSF2001/12)).

Acknowledgments

We acknowledge our university for providing all the necessary facilities.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Duche, H.; Tsegay, A.T.; Tamirat, K.S. Identifying Risk Factors of Breast Cancer Among Women Attending Selected Hospitals of Addis Ababa City: Hospital-Based Unmatched Case-Control Study. *Breast Cancer: Targets and Therapy* **2021**, *13*, 189-197, https://doi.org/10.2147/BCTT.S293867.
2. Moga, M.A.; Dimienescu, O.G.; Bălan, A.; Dima, L.; Toma, S.I.; Bîgiu, N.F.; Blidaru, A. Pharmacological and therapeutic properties of Punica granatum phytochemicals: Possible roles in breast cancer. *Molecules* **2021,** *26*, 1054-1074, https://doi.org/10.3390/molecules26041054.
3. Selvakumar, P.; Badgeley, A.; Murphy, P.; Anwar, H.; Sharma, U.; Lawrence, K.; Lakshmikuttyamma, A. Flavonoids and other polyphenols act as epigenetic modifiers in breast cancer. *Nutrients* **2020**, *12*, 761-778, https://doi.org/10.3390/nu12030761.
4. Harbeck, N.; Penault-Llorca, F.; Cortes, J.; Gnant, M.; Houssami, N.; Poortmans, P.; Ruddy, K.; Tsang, J.; Cardoso, F. Breast cancer. *Nature Reviews Disease Primers* **2019**, *5*, 1-31, https://doi.org/10.1038/s41572-019-0111-2.
5. Ren, J.X.; Gong, Y.; Ling, H.; Hu, X.; Shao, Z.M. Racial/ethnic differences in the outcomes of patients with metastatic breast cancer: contributions of demographic, socioeconomic, tumor and metastatic characteristics. *Breast Cancer Research and Treatment* **2019**, *173*, 225–237, https://doi.org/10.1007/s10549-018-4956-y.
6. Allemani, C.; Weir, H.K.; Carreira, H.; Harewood, R.; Spika, D.; Wang, X.S.; Bannon, F.; Ahn, J.V.; Johnson, C.J.; Bonaventure, A.; Marcos-Gragera, R.; Stiller, C.; Azevedo E Silva, G.; Chen, W.Q.; Ogunbiyi, O.J.; Rachet, B.; Soeberg, M.J.; You, H.; Matsuda, T.; Miles, S.A. Global surveillance of cancer survival 1995-2009: Analysis of individual data for 25 676 887 patients from 279 population-based registries in 67 countries (CONCORD-2). *The Lancet* **2015**, *385*, 977–1010, https://doi.org/10.1016/S0140-6736(14)62038-9.
7. Agarwal, G.; Pradeep, P.V.; Aggarwal, V.; Yip, C.H.; Cheung, P.S.Y. Spectrum of breast cancer in Asian women. *World Journal of Surgery* **2007** *31*, 1031-1040, https://doi.org/10.1007/s00268-005-0585-9.
8. Leong, S.P.L.; Shen, Z.Z.; Liu, T.J.; Agarwal, G.; Tajima, T.; Paik, N.S.; Sandelin, K.; Derossis, A.; Cody, H.; Foulkes, W.D. Is Breast cancer the same disease in Asian and Western countries? *World Journal of*

Surgery **2010**, *34*, 2308-2324, https://doi.org/10.1007/s00268-010-0683-1.
9. Raina, V.; Bhutani, M.; Bedi, R.; Sharma, A.; Deo, S.V.S.; Shukla, N.K.; Mohanti, B.K.; Rath, G.K. Clinical features and prognostic factors of early breast cancer at a major cancer center in North India. *Indian Journal of Cancer* **2005**, *42*, 36-41, https://doi.org/10.4103/0019-509x.15099.
10. Roy, R.; Chun, J.; Powell, S.N. BRCA1 and BRCA2: Different roles in a common pathway of genome protection. In *Nature Reviews Cancer*, **2012**, *12*, 68-78, https://doi.org/10.1038/nrc3181.
11. Prabhavathi, H.; Dasegowda, K.R.; Renukananda, K.H.; Lingaraju, K.; Naika, H.R. Exploration and evaluation of bioactive phytocompounds against BRCA proteins by in silico approach. *Journal of Biomolecular Structure and Dynamics* **2020**, 1-15, https://doi.org/10.1080/07391102.2020.1790424.
12. Shahid, T.; Soroka, J.; Kong, E.H.; Malivert, L.; McIlwraith, M.J.; Pape, T.; West, S.C.; Zhang, X. Structure and mechanism of action of the BRCA2 breast cancer tumor suppressor. *Nature Structural and Molecular Biology* **2014**, *21*, 962-968, https://doi.org/10.1038/nsmb.2899.
13. Lomonosov, M.; Anand, S.; Sangrithi, M.; Davies, R.; Venkitaraman, A.R. Stabilization of stalled DNA replication forks by the BRCA2 breast cancer susceptibility protein. *Genes and Development* **2003**, *17*, 3017-3022, https://doi.org/10.1101/gad.279003.
14. Schlacher, K.; Christ, N.; Siaud, N.; Egashira, A.; Wu, H.; Jasin, M. Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11. *Cell* **2011**, *145*, 529-542, https://doi.org/10.1016/j.cell.2011.03.041.
15. Roy, A.; Bhatia, K.S. In silico analysis of plumbagin against cyclin-dependent kinases receptor. *Vegetos* **2021**, *34*, 50-56, https://doi.org/10.1007/s42535-020-00169-8.
16. Garg, S.; Roy, A. In silico analysis of selected alkaloids against main protease (Mpro) of SARS-CoV-2. *Chemico-Biological Interactions* **2020**, *332*, 109309, https://doi.org/10.1016/j.cbi.2020.109309.
17. Garg, S.; Anand, A.; Lamba, Y.; Roy, A. Molecular docking analysis of selected phytochemicals against SARS-CoV-2 M pro receptor. *Vegetos* **2020**, *33*, 766-781, https://doi.org/10.1007/s42535-020-00162-1.
18. Ram, G.; Sharma, V.R.; Sheikh, I.; Sankhyan, A.; Aggarwal, D.; Sharma, A.K. Anti-cancer potential of natural products: recent trends, scope and relevance. *Letters in Applied NanoBioScience* **2020**, *9*, 902-907, https://doi.org/10.33263/LIANBS91.902907.
19. Agarwal, H.; Malviya, R.; Kumar Sharma, P. Evaluation of in situ gel forming system loaded with etoricoxib and herbal adjuvant nanoparticles against human colon cancer cell lines (HT-29). *Letters in Applied NanoBioScience* **2020**, *9*, 789-795, https://doi.org/10.33263/LIANBS0091.789795.
20. Roy, A. Hairy Root Culture an Alternative for Bioactive Compound Production from Medicinal Plants. *Current Pharmaceutical Biotechnology* **2021**, *22*, 136-149, https://doi.org/10.2174/1389201021666201229110625.
21. Roy, A.; Bharadvaja, N. Venom-derived bioactive compounds as potential anti-cancer agents: a review. *International Journal of Peptide Research and Therapeutics* **2021**, *27*, 129–147, https://doi.org/10.1007/s10989-020-10073-z.
22. Sachdeva, V.; Roy, A.; Bharadvaja, N. Current prospects of nutraceuticals: A review. *Current pharmaceutical biotechnology* **2020**, *21*, 884-896, https://doi.org/10.2174/1389201021666200130113441.
23. Roy, A. Plumbagin: A potential anti-cancer compound. *Mini reviews in medicinal chemistry*, **2021,** *21*, 731-737, https://doi.org/10.2174/1389557520666201116144421.
24. Roy, A.; Bharadvaja, N. Medicinal plants in the management of cancer: a review. *International Journal of Complementary and Alternative Medicine* **2017**, *9*, 00291, https://doi.org/10.15406/ijcam.2017.09.00291.
25. Khatoon, E.; Banik, K.; Harsha, C.; Sailo, B.L.; Thakur, K.K.; Khwairakpam, A.D.; Kunnumakkara, A.B. Phytochemicals in cancer cell chemosensitization: Current knowledge and future perspectives. In *Seminars in Cancer Biology*, **2020,** S1044-579X(20)30150-4, https://doi.org/10.1016/j.semcancer.2020.06.014.
26. Qawoogha, S.S.; Shahiwala, A. Identification of potential anti-cancer phytochemicals against colorectal cancer by structure-based docking studies. *Journal of Receptors and Signal Transduction* **2020**,*40*, 67-76, https://doi.org/10.1080/10799893.2020.1715431.
27. Prabhavathi, H.; Dasegowda, K.R.; Renukananda, K.H.; Lingaraju, K.; Naika, H.R. Exploration and evaluation of bioactive phytocompounds against BRCA proteins by in silico approach. *Journal of Biomolecular Structure and Dynamics* **2020**, 1-15, https://doi.org/10.1080/07391102.2020.1790424.
28. Baseer, S.; Khan, S.; Nouroz, F. Computational drug Cumintrazole-analog for the treatment of mutant BRCA1 and BRCA2 proteins in breast cancer. *International Current Pharmaceutical Journal* **2015**, *4*, 410-414, https://doi.org/10.3329/icpj.v4i8.24021.

PAPER

Fabrication and multi-objective optimization of friction stir processed aluminium based surface composites using Taguchi approach

To cite this article: Ravi Butola *et al* 2021 *Surf. Topogr.: Metrol. Prop.* **9** 025044

View the article online for updates and enhancements.



IOP ebooks™

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection–download the first chapter of every title for free.

This content was downloaded from IP address 14.139.251.100 on 07/07/2021 at 07:33

Surface Topography: Metrology and Properties



PAPER

RECEIVED
14 February 2021

REVISED
29 May 2021

ACCEPTED FOR PUBLICATION
15 June 2021

PUBLISHED
24 June 2021

Fabrication and multi-objective optimization of friction stir processed aluminium based surface composites using Taguchi approach

Ravi Butola[1], **Prakash Chandra**[2], **Kartikeya Bector**[1] **and Ranganath M Singari**[1]

[1] Department of Mechanical Engineering, Delhi Technological University, Delhi-110042, India
[2] Department of Biotechnology, Delhi Technological University, New Delhi-110042, India

E-mail: ravibutola33855@gmail.com

Keywords: Nano-particle, surface composite, optimization, microhardness, residual stress

Abstract

Friction stir processing is a technique for the modification and fabrication of surface composites. Surface composites of aluminium have a diverse application in industries like aerospace and automobile. Here, B_4C nano-particles (< 30 nm) have been employed as the reinforcement on AA7075 based substrate alloy and surface composites have been constructed at varying Tool rotation and Tool traverse speeds. The influence of these varying factors on the ultimate tensile strength (UTS), microhardness (Hv) and residual stress (RS) have been studied. Taguchi L9 orthogonal array was used for the DOE and the changes in the mechanical properties of these nine samples with regards to the tool rotation speeds of 800 rpm, 1000 rpm and 1200 rpm and the tool traverse speeds of 40 mm min^{-1}, 50 mm min^{-1} and 60 mm min^{-1} were investigated. It was found that between the two processing variables, Tool traverse speed was the more significant variable. 800 rpm and 60 mm min^{-1} were determined to be the optimum parameters for Friction stir processing and the predicted values of microhardness (Hv) through Taguchi analysis were the most accurate.

1. Introduction

Friction stir processing (FSP) was pioneered in 2002 and has since then been extensively applied in the alteration of surfaces and fabrication of surface composites. Many other techniques like stir casting [1], chemical vapour deposition [2], powder metallurgy [3], plasma spraying [4], laser machining [5] and suchlike have already been studied extensively and they often depend on the melting of metal for its processing. FSP, on the contrary, focuses on the alteration of material without melting it and is a widely accepted solid processing technique.

FSP has been developed on the fundamental concept of FS Welding and has proliferated as the go-to technique for the processing of substrates of Al, Cu and Mg with the reinforcements like B_4C, SiC, Al_2O_3 and some natural reinforcements as well. Further, research has been conducted on the use of nano-particles as reinforcement in these substrates, improving the chemical stability and the microstructural profile of the hence fabricated surface composites. For the application of reinforcements on the surface, many methods have been explored, the prominent ones comprising of groove, holes and use of volatile media mixed with the particles for smearing.

Given the resulting refinement of microstructure, FSP has been accepted as an industry standard in aerospace and automobile; in addition to the surface refinement of the aforementioned metal substrates, it is also being used on steel, titanium and so forth [6]. Chaudhary *et al* [7] reviewed the efficacy of FSP in surface refinement and how this has also lead to the enhancement of the mechanical and tribological properties of surface composites of various alloys. Butola *et al* [8] scrutinized the hardness and corrosion behaviour, amongst other features, of the MMCs formed by natural reinforcements. B_4C is a hard particle and its incorporation into the metal matrix has been shown to improve the mechanical properties of the specimen manifold. Yuvaraj *et al* [9] examined the mechanical and tribological properties of Al5083/B_4C FSPed SC as a function of number of passes and reinforcement size. Butola *et al* [10] studied the effect of nano-particles of B_4C on AA7075-T6 substrate when the self-assembled monolayer technique was employed to produce surface composites via FSP. They investigated the tensile strength, microhardness

© 2021 IOP Publishing Ltd

and other microstructural properties of the SC. Wang *et al* [11] manufactured AZ31 Mg/SiC_p reinforced SCs and studied their microhardness and wear properties. Ramezani *et al* [12] studied the effect of rotation speed, traverse speed and number of passes on the tool wear, microhardness and topography of FSPed AA7075/SiC nanocomposites through RSM technique. They concluded that traverse speed was the most significant variable, followed by rotation speed and pass number respectively. Bauri *et al* [13] investigated the effect of process parameters viz rotational speed and traverse speed, and tool geometry, on the incorporation of the Ni particles in the Al 5083 alloy. They found that FSP at 1200 rpm and 0.4 mm s^{-1} were the minimum requirements to successfully introduce the Ni particles in the metal matrix. They also observed a more uniform distribution of the particles in the matrix, after FSP, along with better hardness and tensile properties. Węglowski *et al* [14] used the trepanation method to study the residual stress in the FSPed AlSi9Mg and concluded that it progresses from the retreating to the advancing end of processing. Mehdi and Mishra [15] used the portable Pulstec XRD machine to investigate the residual stress in the nugget area of TIG + FSP join of dissimilar aluminium alloys. Butola *et al* [16] measured the residual stress on the H13 tool steel used to fabricate tool pins for FSP.

The surface composites properties are heavily dependent on the processing parameters of FSP. It, hence, becomes crucial for one to not only tweak the right variables, but also understand the optimum parameters for the composite to have certain desired properties. Various techniques have been adopted for this optimization. Taguchi, RSM and Grey-Fuzzy include a bulk of them.

Rathee *et al* [17] optimized the process parameters through Taguchi L9 orthogonal array to achieve the maximum hardness of AA6061/SiC surface composite formed by FSP. It was observed that hardness is maximum when tool rotational speed, transverse speed and tilt angle are at 1400 rpm, 50 mm min^{-1} and 2.5 deg. Ahmadkhaniha *et al* [18] determined the best combination of process parameters that influence the hardness of FSPed pure Mg by using Taguchi optimization and through analysis of variance it was found that rotational speed, transverse speed and tool tilt angle are the most significant parameters that affects the hardness of the sample. Butola *et al* [19] reviewed the use of numerous approaches like Taguchi, RSM and Grey-Fuzzy Algorithm to optimize CNC turning parameters. Bector *et al* [20] reviewed using ANN for the modelling of FSP. Ahmad *et al* [21] optimized the process variables of MAF through Taguchi technique. Butola *et al* [22] employed Taguchi to optimize the processing variables of FSP for the fabrication of AA7075 surface composites with SiC, RHA and B_4C. They observed that in the stir zone, now smaller in size, the most uniform dispersion of reinforcement was realized with a square mandrel. Butola *et al* [23] also optimized abrasive flow machining through the Taguchi method.

This research work aims to incorporate B_4C nanoparticles with diethyl ether (DEE) solution over the groove surface of AA7075 based matrix using Micropipette, followed by Friction stir processing. Further, we have determined the impact of changing tool rotation speed and tool traverse speed on the tensile strength and microhardness of the thus fabricated FSPed surface composites.

2. Materials and methods

2.1. Materials

Aluminium alloy 7075-T6 was used as the substrate, with nano-particles of B_4C (size$<$30 nm) as the reinforcement. The experimental tests were performed on friction stir welding setup and the dimension of the workpiece was decided with regards to the FSP set up (200 × 80 × 6 mm). A groove of dimensions 160 mm length, 2 mm width and 2 mm depth was made and the B_4C nano-particle blended with die ethyl ether (DEE), was filled in it-shown in figure 1(b). The groove was processed using a square pin tool of H13 tool steel having 5.10 mm pin side with 3 mm pin length and 19.95 mm shoulder diameter.

2.2. Design of experiment

The experiment was designed on Minitab 17 software using Taguchi L9 Orthogonal Design of Experiment (DOE) is shown in Table 1. The tool rotation speed and tool traverse speed were the chosen process parameters. A 2-factor 3-level DOE was done by taking these variables into consideration as mentioned in Table 2. The other FSP parameters such as tool pin profile (square), number of passes (1) and tilt angle (0° in the covering pass to cover the groove and 2° during composite fabrication process) were taken as constant. For the covering of the reinforced grooves, a pinless tool was used.

2.3. Fabrication of surface composite

FSP is a relatively new technique that can be utilized for surface modification and manufacturing of SCs. FSP involves a high speed, rotating, non-consumable tool that slides along a workpiece gradually. The interaction of the shoulder of the tool with the workpiece surface, while a pivotal power is applied to it, causes plastic deformation. The rotating tool is then moved along the workpiece in the processed direction. Considerable heat is produced because of traction amidst the tool shoulder and the specimen. The hot, softened material moves along the processing line and below the back-up rim to the tool's end, where it's compacted and combined through critical distortion, following which it hardens. The arrangement of the schematic diagram of FSP, along with the apparatus, is shown in figures 1(a) and (b) respectively. B_4C was considered as the reinforcement material and a square

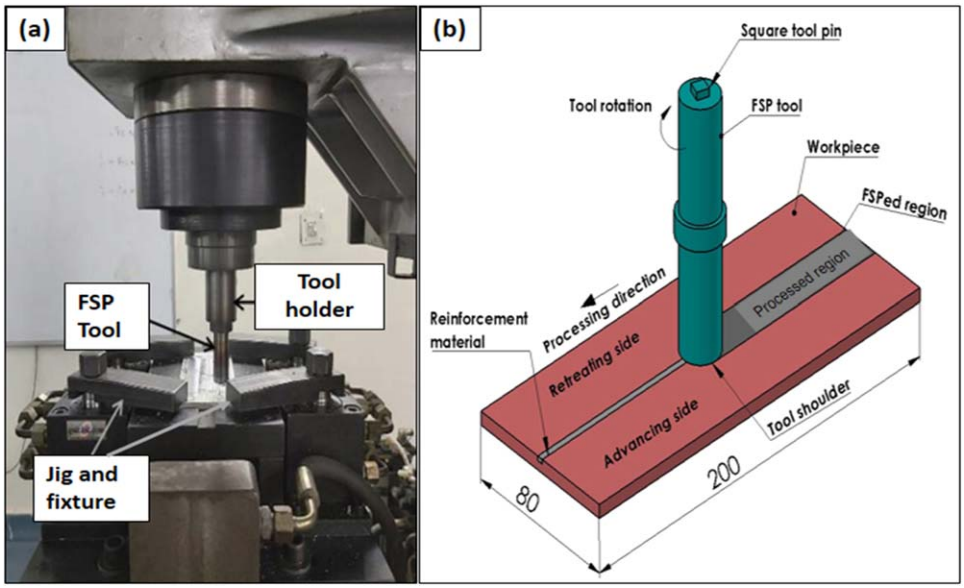


Figure 1. (a) FSW/FSP Apparatus (b) Schematic diagram of FSP.

Table 1. DOE using taguchi L9 orthogonal array.

Sample number	Tool rotation speed	Tool traverse speed
S1	800	40
S2	1000	40
S3	1200	40
S4	800	50
S5	1000	50
S6	1200	50
S7	800	60
S8	1000	60
S9	1200	60

Table 2. Selection of process parameters.

Processing Parameter	Unit	Level 1	Level 2	Level 3
Tool Rotation Speed	RPM	800	1000	1200
Tool Traverse Speed	mm min^{-1}	40	50	60

tool pin profile was used for the fabrication of the surface composite using groove method [22].

It crucial for one to consider the physical effects of tool pin profiles at the tool-sample interface [24]. The geometry and size of the tool pin considerable changes the amount of heat engendered and the material flow in FSP. A tool with a big shoulder diameter leads to the concentration of the frictional heat. Frictional heat is created at the points where the rotating tool shoulder presses onto the workpiece, which further plasticizes the metal and pushed it downwards. The frictional heat engendered by the longitudinal and rotational movement of the pin and due to the tool shoulder softens the metal at positions right adjacent to the pin. The metal melds round the pin, joining the weld seam. This leads to much better refined particles in the 2° phase and, consequently, a more stable microstructure. Many studies [25–27] have shown that amid various pin profiles, the square pin profile is the most effective. Hence, the surface composite was fabricated using a square pin profile (figure 1(b)).

2.4. Ultimate tensile strength testing

Wire EDM was used for the extraction of bone-shaped samples to test the ultimate tensile strength (UTS) of the processed samples, in accordance with ASTM: E8 standard. The measurement of the UTS was done through the computer-controlled universal testing machine. The specimens were held from the ends and an increasing load was applied throughout the test, till the breaking point was obtained. The changes in the samples were recorded and then analysed.

2.5. Microhardness testing

For microhardness testing, the Fischer scope HM2000S was used. The testing was done according to ASTM-E8 standard. To measure the Vickers hardness, a diamond indenter was pushed on the surface of the specimen fixed on the indentation lens. The force was incremented from 300 mN to 500 mN for a creep time of 5 s. The operating time was around 40 s and the Hv property was for standard application of hardness from the properties taskbar. After the zero point was identified by the machine, the force was increased from 0 to 500 mN for 20 s. A creep time of 5 s was observed and then the force was decreased back to 0 in 20 s. The readings were shown in a tabular and graphical form for n variable values, where n is the number of readings for a specific specimen. Four readings were recorded for each specimen and the average Vickers hardness was noted.

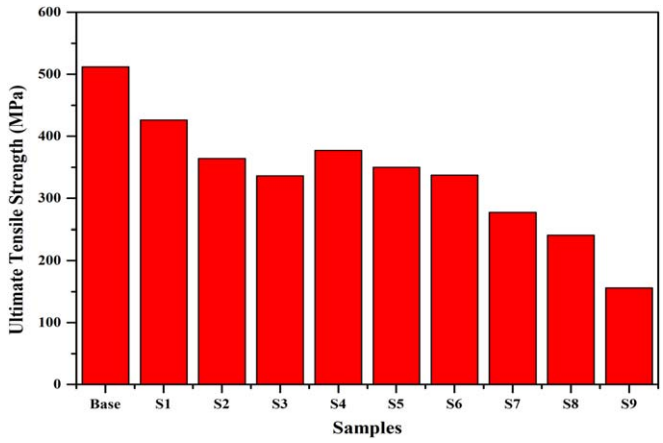


Figure 2. Experimental values of ultimate tensile strength.

Table 3. Experimental values of all specimens.

Sample number	Tool rotation speed	Tool traverse speed	Ultimate tensile strength (MPa)	Microhardness (Hv)	Residual stress	Residual stress^{-1}*10^2 (scaled)
S1	800	40	426	175	17.67	5.67
S2	1000	40	364	181	27.34	3.66
S3	1200	40	336	193	29.34	3.41
S4	800	50	377	169	15.34	6.52
S5	1000	50	350	172	17	5.88
S6	1200	50	337	187	22.67	4.41
S7	800	60	277	166	12.67	7.89
S8	1000	60	241	171	16.67	5.99
S9	1200	60	156	176	23.34	4.29

2.6. Residual stress testing

Residual stresses are the forces remaining in the processed surface even after the removal of external load, or the FSP tool in this case. These significantly affect the fatigue behaviour of the sample and hence impact its life. Conventional residual stress measurement techniques such as numerical simulation have failed to provide accurate values in spots of miniscule diameters. This has led to the adoption of modern techniques like x-ray diffraction (XRD)—which is a non-destructive method. It is primarily based on the Bragg's equation—a function of interplanar distance D_o and the diffraction angle η_o. Two major types of XRD techniques exist—$\sin^2$ and cos methods. In recent times, cos method has emerged as the more widely accepted method. As the sample is blasted with x-rays, the crystalline structure of the sample diffracts the x-rays into different directions. For each of these diffractions, a diffraction peak can be observed for stress computation. This method is leads to the formation of Debye rings and unlike the $\sin^2$ method, there's no need for the sample to be irradiated at varied angles. This Debye-ring can be compared to a standard one for the calculation of stress. In this study, we have used a portable machine—Pulstec Stress Analyzer α360 Portable Stress Analyzer—for the measurement of residual stress using the cos method.

3. Results and discussion

It can be seen from table 3 and figure 2 that an increment in the tool rotation speed led to the UTS value of the specimens decreasing due to the increment in the microhardness value. Further, with an increase of the tool traverse speed, the UTS value again decreased. In all, an increment in either of the parameters led to a decrease in the UTS values of the specimens. The highest UTS of fabricated specimens was observed for S1 while the lowest UTS was for S9.

For microhardness, as can be observed from table 3 and figure 3, increasing the tool rotation speed led to an increment in the microhardness (Hv) value. Asadi *et al* [27] and Butola *et al* [28] had similar observations. It may also be generalized that increasing the tool traverse speed leads to the microhardness (Hv) value being decreased. The highest microhardness value was seen for S3 and the lowest for S7. The value of microhardness for the base alloy was 102 Hv. Mishra *et al* [29] and Paidar *et al* [30] also observed that increasing the tool traverse speed decreased the microhardness of the FSPed sample due to poor distribution of the particles in the stir zone.

For residual stress, as can be observed from table 3 and figure 4, increasing the tool rotation speed led to an increment in the residual stress value. Further, an

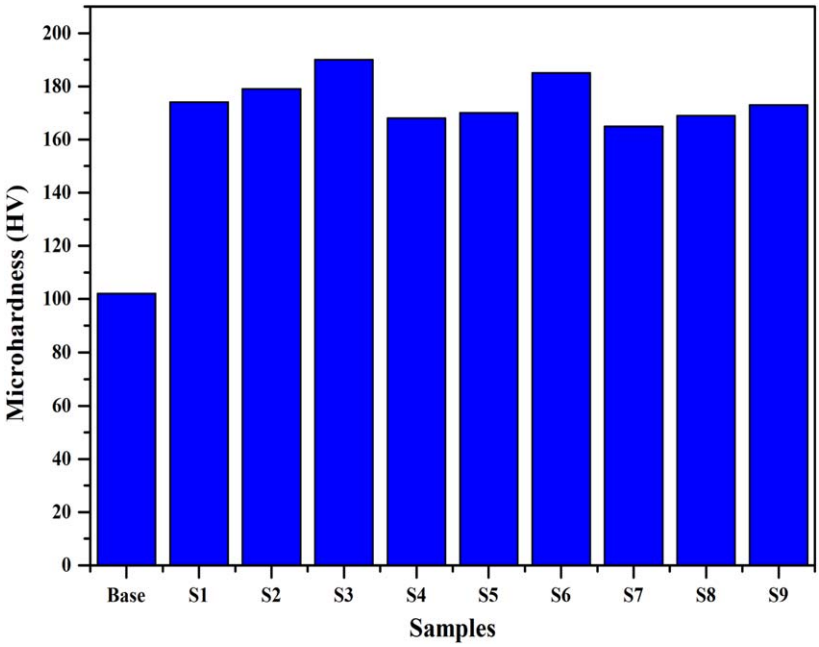


Figure 3. Experimental values of microhardness.

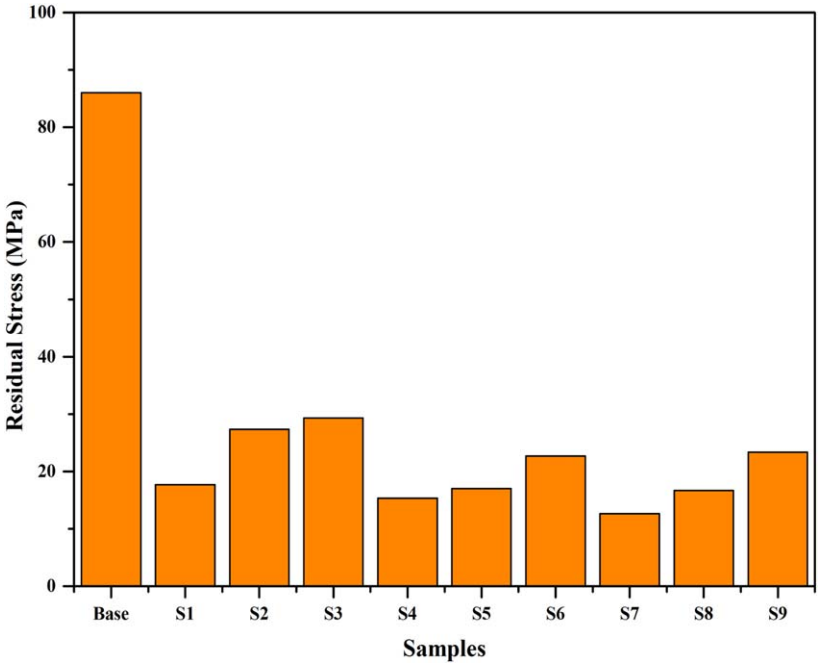


Figure 4. Experimental values of residual stress.

increase in the tool traverse speed resulted in the reduction in the residual stress values. The highest residual stress value was seen for S3 and the lowest for S7. The value of residual stress for the base alloy was 86 MPa.

3.1. Microstructural analysis

For the observation of microstructure, an inverted metallurgical optical microscope (Olympus GX41) was utilized and the microstructure images were captured at 50 μm. The width of the micrograph was measured as per the pixel to scale conversion scale of 2 pixels per micron.

Figure 5 portrays the microstructural images of the base alloy and the surface composite FSPed at 800 rpm and 1200 rpm. It can be clearly observed that the grain size of the base alloy (figure 5(a)) is larger than that of the surface composites. Furthermore, a higher tool rotation speed of 1200 rpm and traverse speed of 60 mm min^{-1} has ensured finer dispersion of the reinforcements

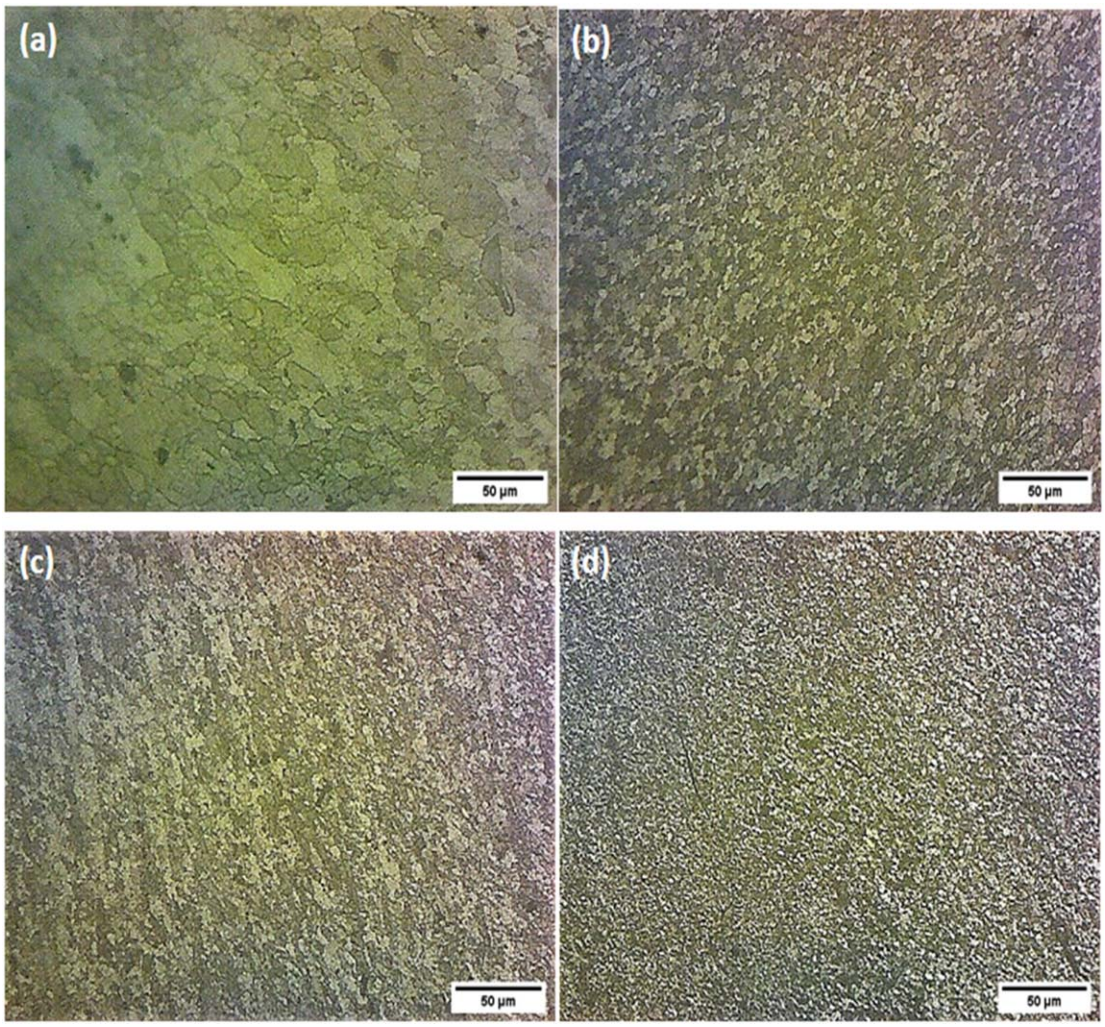


Figure 5. Microstructure of (a) Base alloy (b) SC specimen FSPed at 800 rpm and 40 mm min^{-1} (c) SC specimen FSPed at 1000 rpm and 50 mm min^{-1} (d) SC specimen FSPed at 1200 rpm and 60 mm min^{-1}.

(figure 5(d)) than the composite FSPed at 800 rpm and traverse speed of 40 mm min^{-1} (figure 5(b)) and the composite FSPed at 1000 rpm and traverse speed of 50 mm min^{-1}. A general refinement in the microstructure can also be observed as compared to the base alloy [28]. It has been seen time and again that an increase in tool rotation speed, up to 1200 rpm, and increasing the traverse speed leads to an improvement in the dispersal of reinforcement in the surface of the workpiece. In terms of the tool traverse speed, after a certain value, the tool does not heat up the material sufficiently and hence the improvement in microstructure decreases.

3.2. SEM and EDS analysis

Following the fabrication of the surface composite using FSP, the specimens were cut orthogonally with regards to the FSPed region and then polished using varying grades of emery paper, followed by wet polishing techniques. Field Emission Scanning Electron Microscope (FESEM) model (ZeissSigma300) was used to observe the dispersion of reinforcement particles. Figures 6(a), (b) portrays the SEM micrographs and EDS profiles of the base alloy and surface composite. The distribution of B_4C nano-particles can be seen in the FSPed region in the SEM micrograph.

Figure 7(a) shows the EDS mapping of the base alloy, it reveal the major composition of the base alloy while indicating the absence of the reinforcement in it. Figure 7(b) shows the EDS mapping of the surface composite, it reveal the composition of the fabricated surface composites and indicate a significant presence of boron and carbon particles in the composite, confirming the successful incorporation of the reinforcement viz B_4C in the metal matrix.

Segal [31] reviewed various modes and processes of severe plastic deformation (SPD) and some of the conclusions made by him are applicable in the explanation of the impact of process parameters of FSP on the hardness, tensile strength and residual stress of the SCs as well. There exists a trade-off between the tool rotation speed and the tool traverse speed. As the rotation speed increases, the working temperature, strain rate and the equivalent strain also increases. On the other hand, an increase in the traverse speed leads to a reduction in the equivalent strain while increasing the strain rate. A higher strain speeds up the formation of sub-grain through SPD because of the

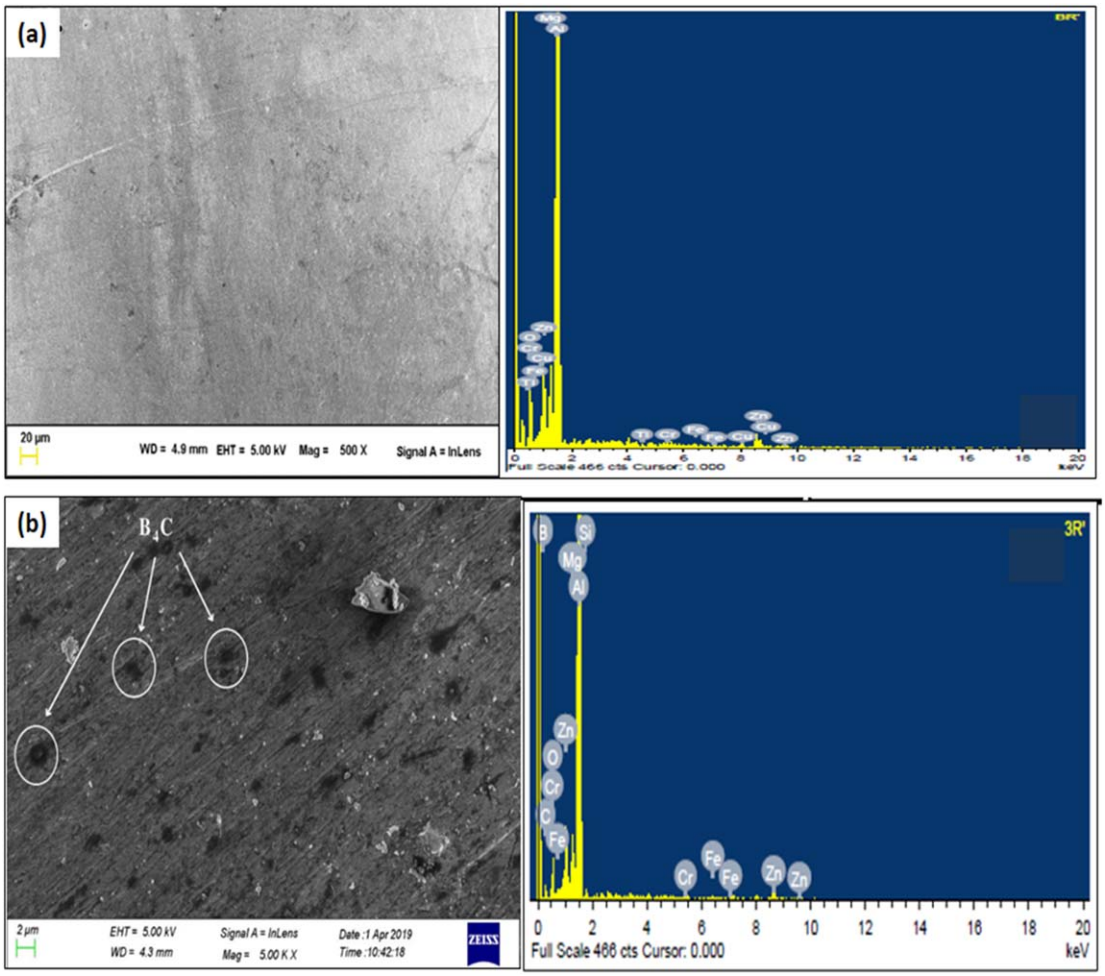


Figure 6. SEM micrograph of (a) base alloy with EDS profile (b) FSPed surface composite with EDS profile.

Table 4. ANOVA of ultimate tensile strength (UTS).

Source	DF	Adj SS	Adj MS	F-Value	P-Value	Significant value
RPM	2	10500	5250.1	9.15	0.032	0.1988
TTS	2	40028	20013.8	34.87	0.003	0.7577
Error	4	2296	573.9			
Total	8	52824				

higher number of dislocation entanglement (akin to the impact of a higher strain rate). It has further been shown that higher rotation speed results in composites having a higher microhardness value (following the application of a higher applied strain at high rotation speed); but tool traverse speed reduces the microhardness due to lower dislocation entanglement because of the higher applied strain reduction. Similar types of results supported by Tyagi *et al* [32].

In the case of residual stress, different types of strain fields exist in the front and back of the tool pin—extrusion-like in front of it and forging behind it. The extrusion-like strain field is tensile, which can increase the residual stress, while the forging stress is compressive and balances out the overall residual stress value. Higher rotation speeds increase the tensile residual stress and hence the overall residual stress value. But in some cases, like in S7, the compressive strain field reduces the overall residual stress value of the specimen.

3.3. Analysis of variance (ANOVA) of UTS

Regression equation

$$\begin{aligned} \text{UTS} = {} & 318.22 + 41.8 * \textit{Tool Rotational Speed}_{800} \\ & + 0.1 * \textit{Tool Rotational Speed}_{1000} \\ & - 57.72 * \textit{Tool Rotational Speed}_{1200} \\ & + 57.1 * \textit{Tool Traverse Speed}_{40} \\ & + 36.4 * \textit{Tool Traverse Speed}_{50} \\ & - 98.6 * \textit{Tool Traverse Speed}_{60} \end{aligned} \quad (1)$$

A close scrutiny of the data in table 4 shows that of the two processing parameters, the Tool Traverse Speed

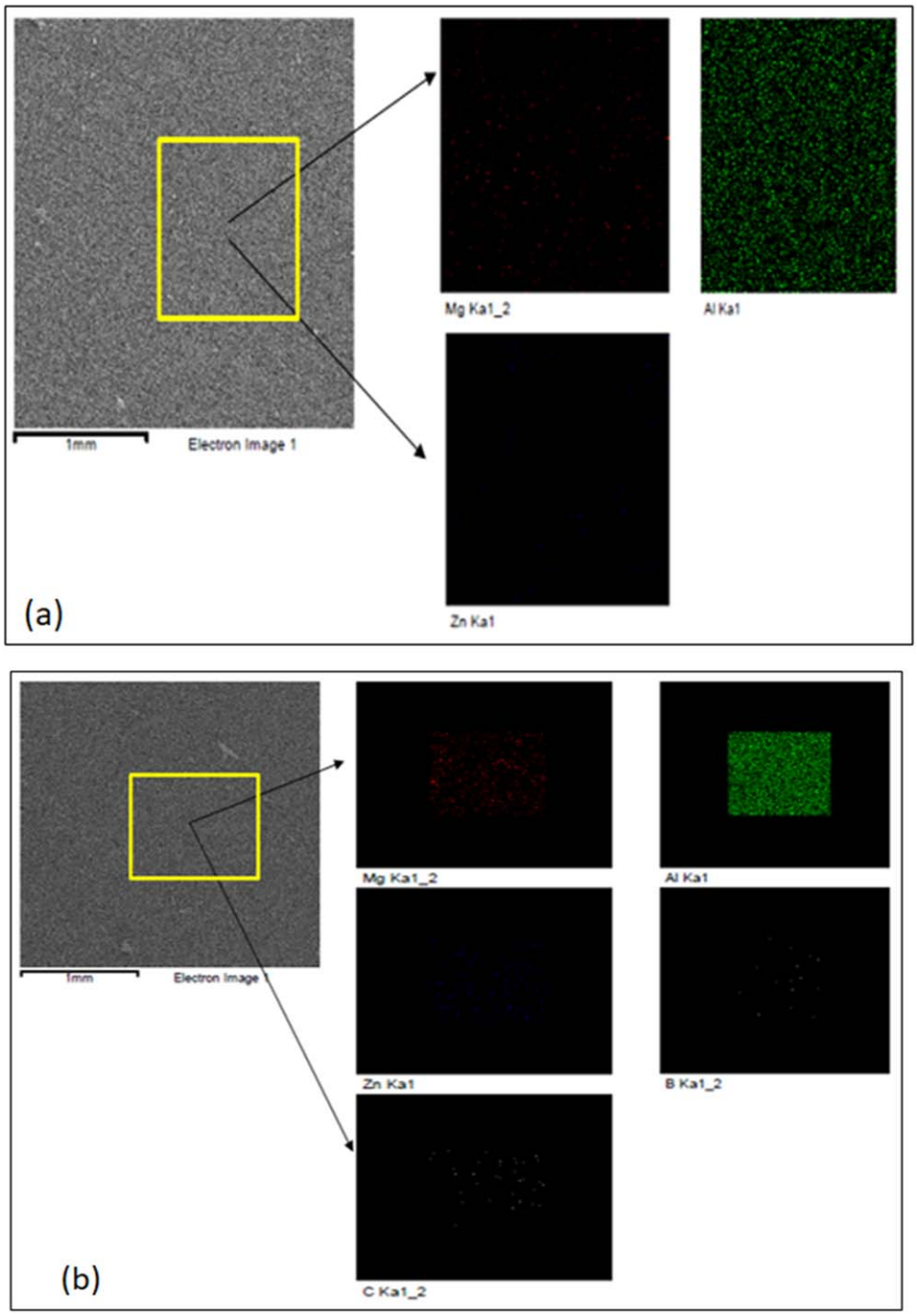


Figure 7. EDS mapping of (a) base alloy and (b) FSPed surface composite.

is more significant. The P-value is a statistical parameter that helps determine the significance of parameters in such a situation. A P-value less than 5% or 0.05 indicated that a variable is significant. Both of the variables can be observed to have a P-value of less than 5% here, but tool traverse speed is more significant in case of UTS. Figure 8 shows the corresponding residual plots having satisfactory plots of normal probability, confirming that there were minimum outliers and that the residuals have a generally constant variance [33]. Equation (1) shows the regression equation for UTS.

3.4. Analysis of variance (ANOVA) of microhardness

Regression equation

$$\begin{aligned} HV = {} & 176.67 - 6.67 * \textit{Tool Rotational Speed}_{800} \\ & - 2.00 * \textit{Tool Rotational Speed}_{1000} \\ & + 8.67 * \textit{Tool Rotational Speed}_{1200} \\ & + 6.33 * \textit{Tool Traverse Speed}_{40} \\ & - 0.67 * \textit{Tool Traverse Speed}_{50} \\ & - 5.67 * \textit{Tool Traverse Speed}_{60} \end{aligned} \quad (2)$$

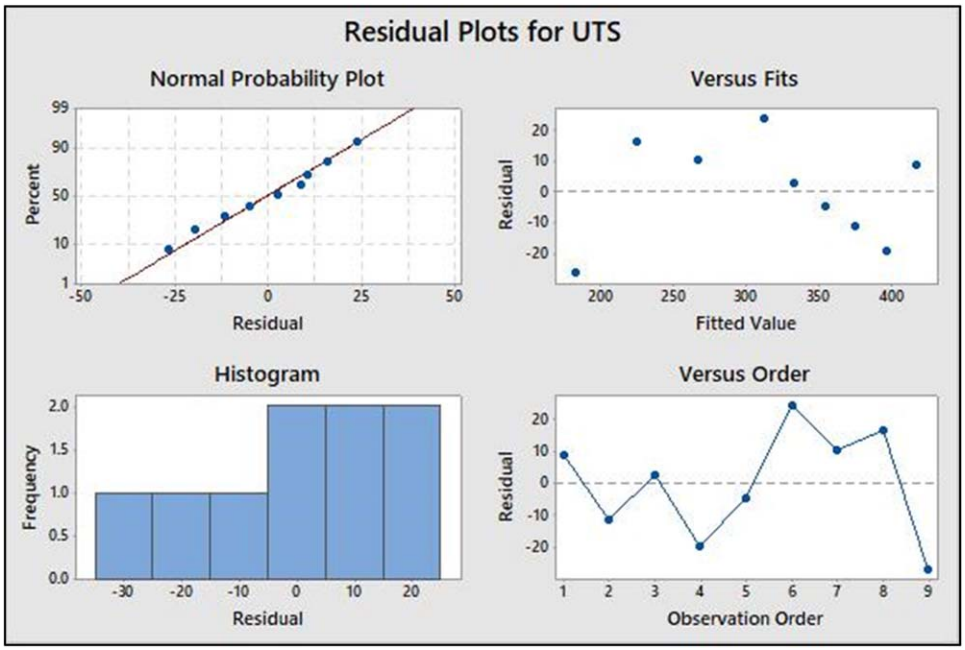


Figure 8. Residual plots for ultimate tensile strength.

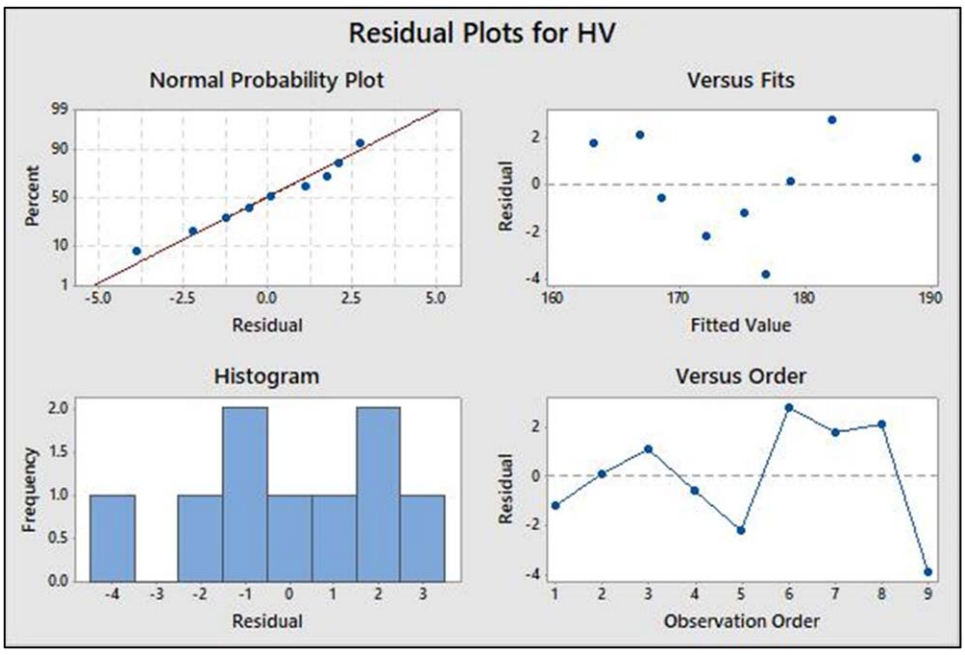


Figure 9. Residual plots for microhardness.

Table 5. ANOVA of microhardness (Hv).

Source	DF	Adj SS	Adj MS	F-Value	P-Value	Significant value
RPM	2	370.67	185.33	22.24	0.007	0.5959
TTS	2	218.00	109.00	13.08	0.018	0.3504
Error	4	33.33	8.33			
Total	8	622.00				

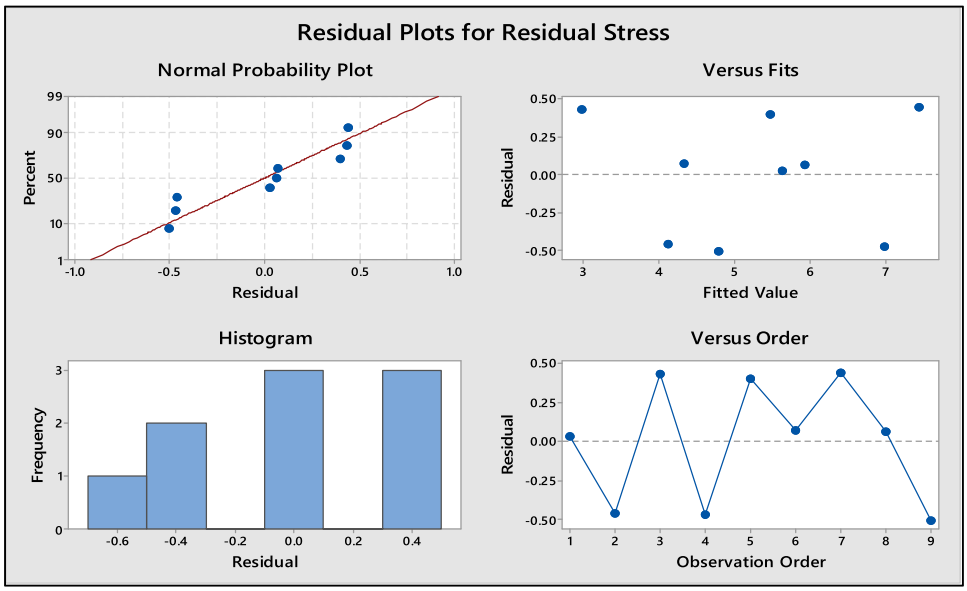


Figure 10. Residual plots for residual stress.

Table 6. ANOVA of residual stress.

Source	DF	Adj SS	Adj MS	F-Value	P-Value	Significant value
RPM	2	10.652	5.3260	17.10	0.011	0.617
TTS	2	5.364	2.6818	8.61	0.036	0.311
Error	4	1.246	0.3115			
Total	8	17.262				

Table 5 has the ANOVA for HV and in this case, the Tool Rotation Speed is more significant. Again, for both parameters, the P-value is less than 5%, but the tool rotation speed holds more significance than tool traverse speed for microhardness. Figure 9 shows the corresponding residual plots having satisfactory plots of normal probability, confirming that there were minimum outliers and that the residuals have a generally constant variance. Equation (2) shows the regression equation for microhardness.

3.5. Analysis of variance (ANOVA) of residual stress

Regression equation

$$\begin{aligned} \boldsymbol{RS} = {} & 5.30 + 1.39^{*}\boldsymbol{Tool\ Rotational\ Speed}_{800} \\ & - 0.12 * \boldsymbol{Tool\ Rotational\ Speed}_{1000} \\ & - 1.26 * \boldsymbol{Tool\ Rotational\ Speed}_{1200} \\ & - 1.06 * \boldsymbol{Tool\ Traverse\ Speed}_{40} \\ & + 0.30 * \boldsymbol{Tool\ Traverse\ Speed}_{50} \\ & + 0.76 * \boldsymbol{Tool\ Traverse\ Speed}_{60} \end{aligned} \quad (3)$$

Table 6 shows the ANOVA for residual stress and here, the Tool traverse speed holds more significance. Here, the P-value is less than 5% for none of the parameters, but the tool traverse speed is more significant than tool rotation speed for residual stress. Figure 10 shows the corresponding residual plots having satisfactory plots of normal probability, confirming that there were minimum outliers and that the residuals have a generally constant variance. Equation (3) shows the regression equation for residual stress.

Table 7. SN ratio for multi-objective taguchi analysis.

Level	Tool rotation speed	Tool traverse speed
1	21.19	17.09
2	18.84	19.62
3	16.83	20.14
Delta	4.36	3.05
Rank	1	2

3.6. Multi-objective taguchi analysis

The Taguchi Analysis aims to provide a more holistic view of the influence of the processing variables on both response variables simultaneously. This helps in the determining of the optimum processing parameters. The signal-to-noise ratio (SNR) is a marker of the impact of noise (deviations) on the experimental values. Here, in both UTS and microhardness, the larger-the-better parameter has been used. It may be expressed as equation (4), where n is the number of iterations and Y_i is the performance characteristic for the test in the ith iteration.

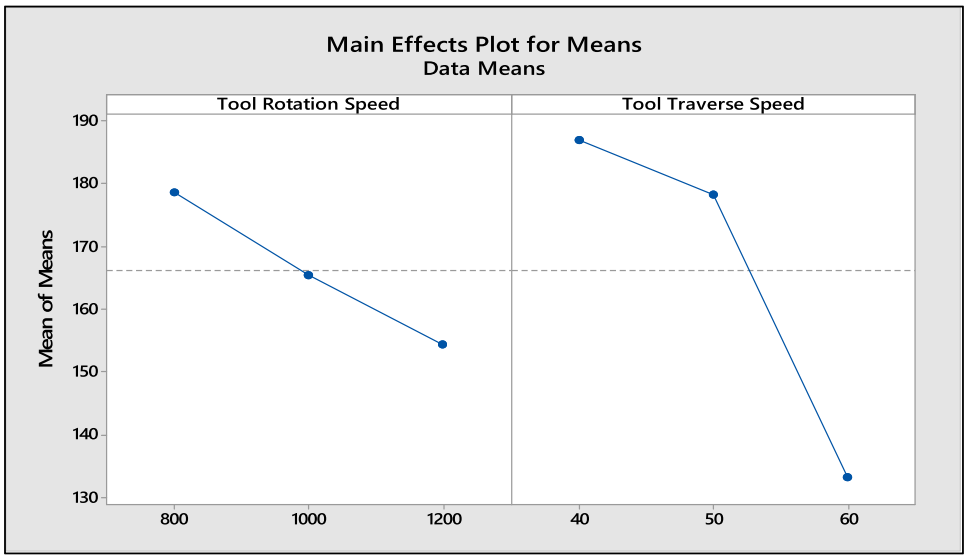


Figure 11. Main effects plot for means.

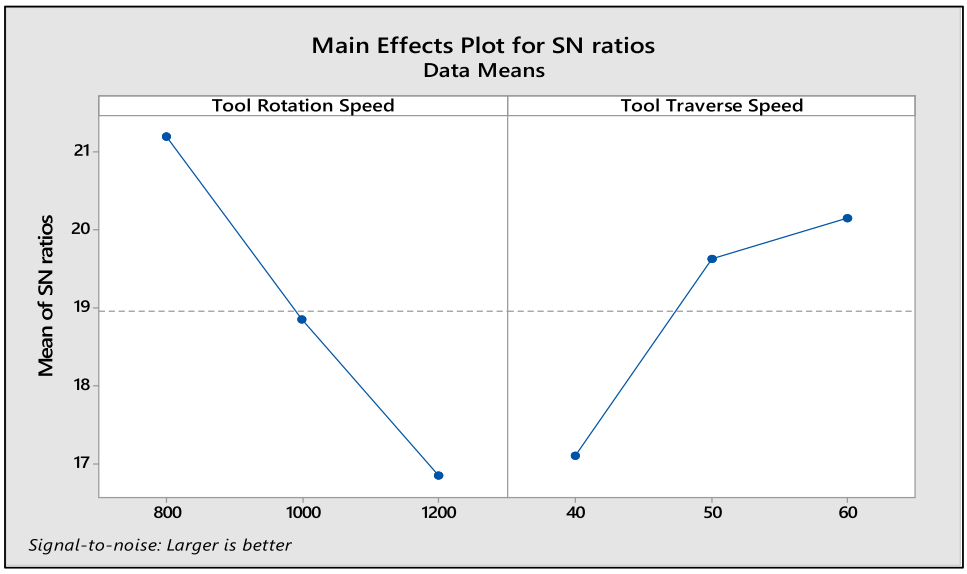


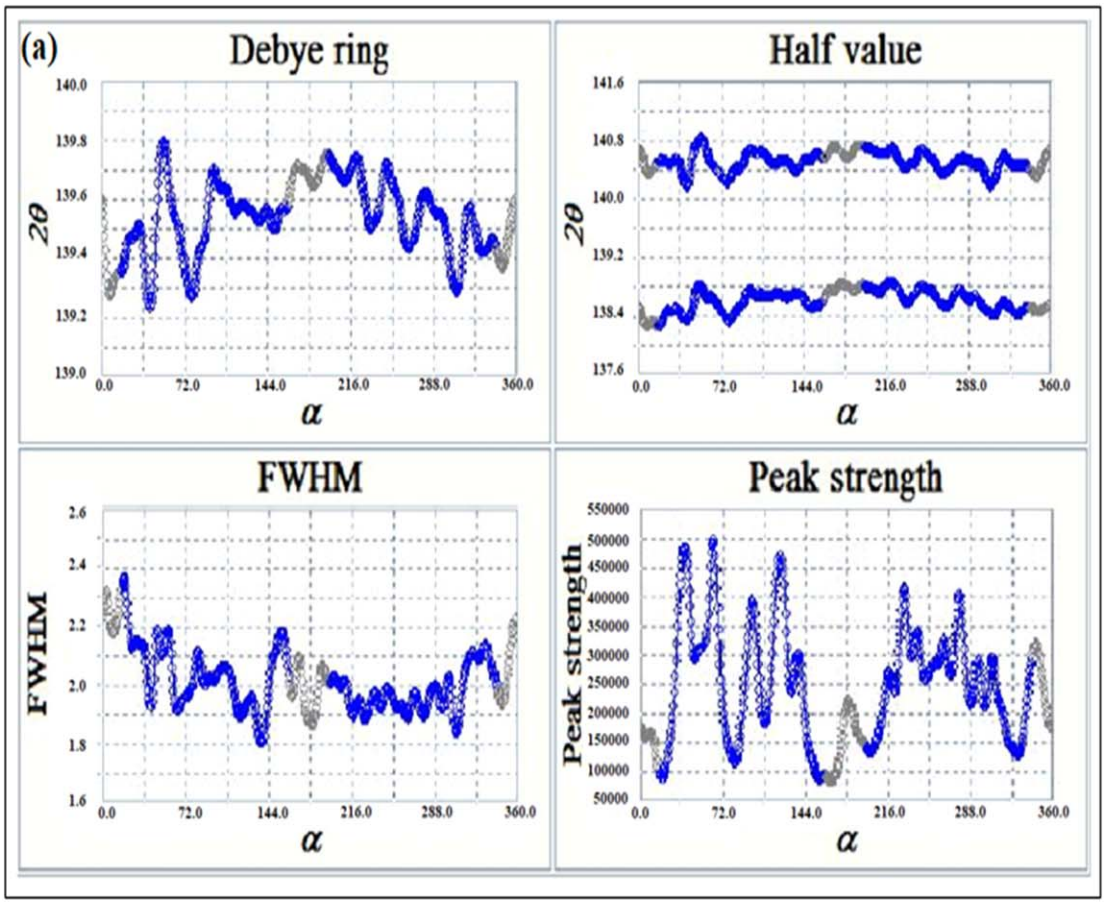
Figure 12. Main effects plot for SNR.

$$\frac{S}{N} = -10 \times \log\left(\frac{1}{n}\sum_{n}^{i=1}\frac{1}{Y_i^2}\right) \quad (4)$$

The main effects plots for averages show the change in the response values with the process variables. The horizontal dashed line is an indicator of the overall average value of the response.

Table 7 shows the response table for SNR for multi-objective analysis. The ranks show that tool traverse speed is the dominant variable. Figure 11 shows the means plot and figure 12 shows the SNR plot. It is evident that the optimal processing values are 800 rpm and 60 mm min^{-1}—S7.

The residual stress of the various specimens was done using the portable x-ray diffraction machine. Here, the residual stresses are calculated using the cosα method. Figures 13 and 14 show the residual stress graphs of the base alloy and sample S7. The peak strength of the base alloy is higher. The linear form of the Debye ring shows the distinct peaks with regards to the different orientations [34]. The Debye ring peaks are greater for the base alloy as well. The Debye rings help us understand the microstructural features of the specimen [35]. The full width at half maximum (FWHM) distribution is measure of surface integrity [36]. A higher FWHM signifies higher residual stresses in the surface. High values of FWHM indicated smaller grain size (and better hardness properties).



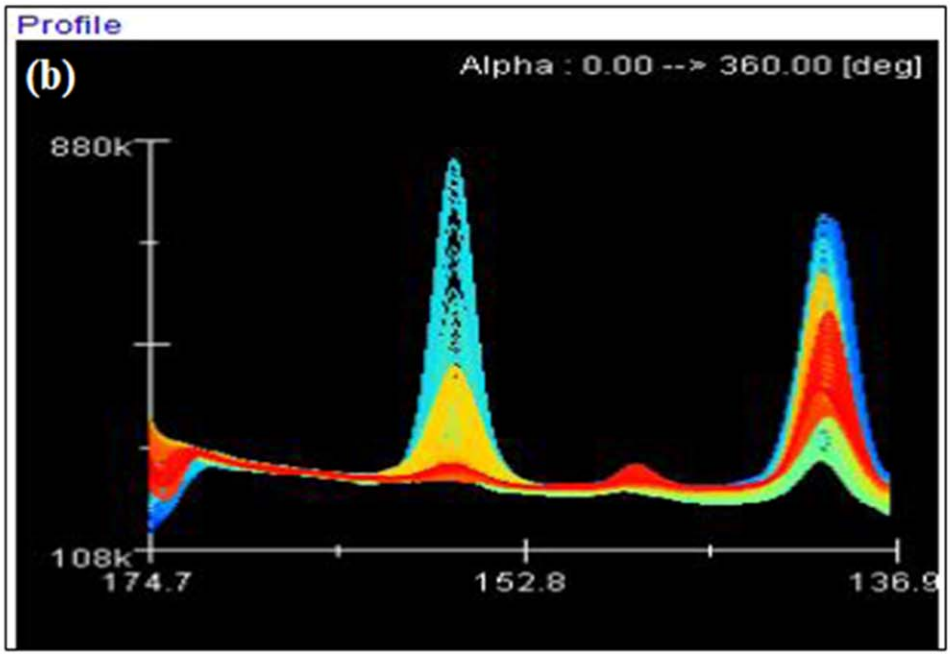


Figure 13. (a) Residual Stress graphs and (b) profile of residual stress of base alloy.

3.7. Accuracy of predictions

Using equations (1)–(3), the values of the response variables have been predicted in table 8. Of the three responses, the microhardness has been predicted with the highest accuracy, followed by UTS and residual stress respectively.

4. Conclusions

In this study, we have investigated the impact of tool rotation speed and tool traverse speed on the friction stir processed surface composite samples by measuring the response variables of ultimate tensile strength,

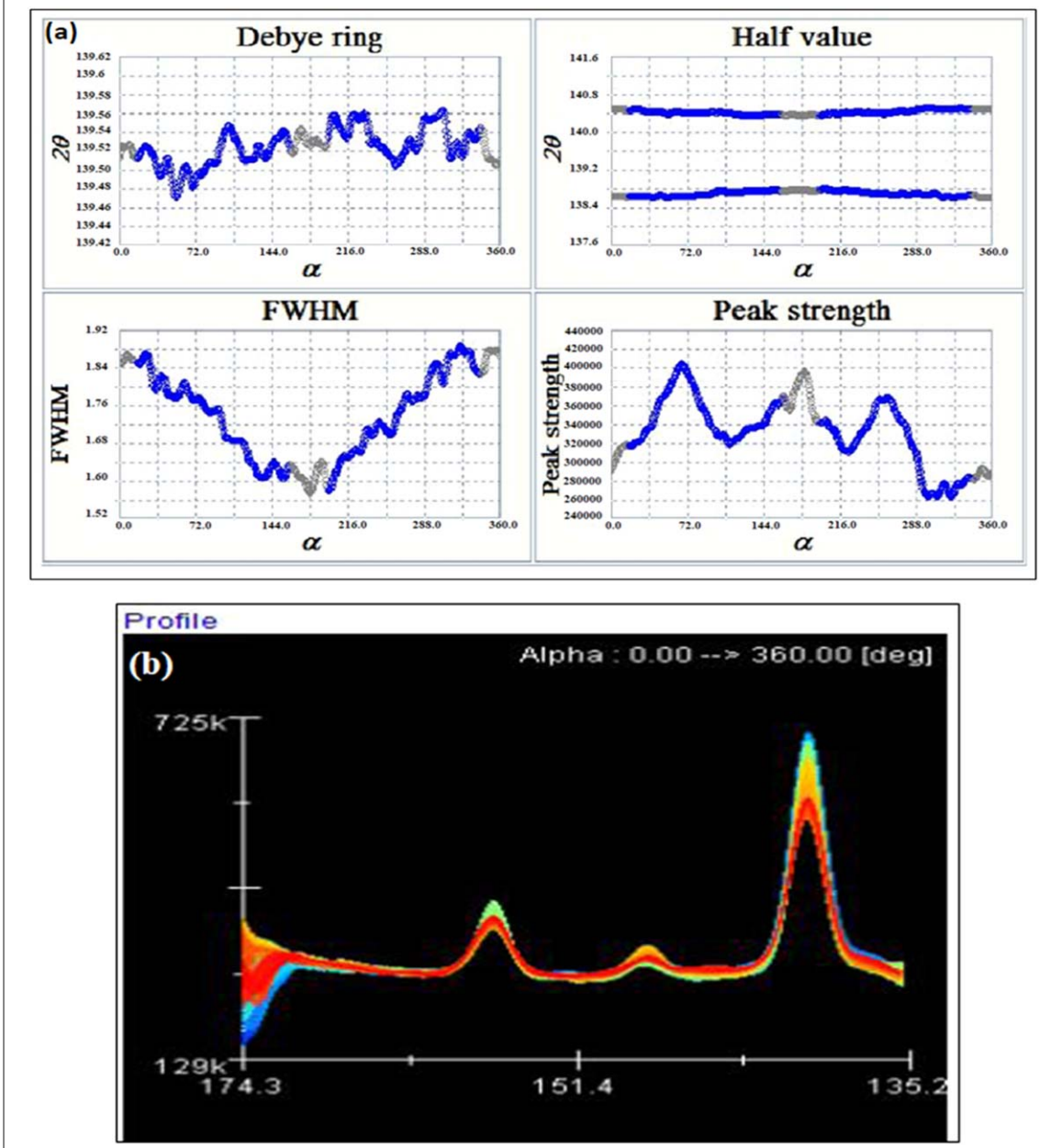


Figure 14. (a) Residual stress graphs and (b) profile of residual stress of S7.

Table 8. Prediction of properties of optimum sample.

Property	Experimental value	Predicted value	% Error
UTS	277	261.42	5.63
HV	166	164.32	1.01
RS	12.67	13.42	5.61

microhardness and residual stress. The points given henceforth can be concluded from the study –

- The Aluminium alloys 7075 based matrix and B_4C nano-particle were incorporated and successfully fabricated surface composites using Friction stir processing.
- Of the given processing parameters, overall, tool traverse speed has been shown to have a higher impact on the mechanical properties than tool rotation speed.
- The optimum parameters were obtained using Taguchi analysis and sample S7 processed at 800 rpm and 60 mm min^{-1} was adjudged to be the optimally processed specimen.
- While the movement of the tool results in different types of residual stress to be generated in front of or behind the tool, the overall residual stress for all FSPed samples was compressive and the least for the optimal sample, S7.
- Predictions made using the Taguchi-generated regression equations proved to be the most accurate for microhardness, followed by UTS and residual stress respectively.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

Ravi Butola https://orcid.org/0000-0002-6581-4131
Prakash Chandra https://orcid.org/0000-0002-8736-7993
Ranganath M Singari https://orcid.org/0000-0003-4981-3857

References

[1] Butola R, Pratap C, Shukla A and Walia R S 2019 Effect on the mechanical properties of aluminum-based hybrid metal matrix composite using stir casting method *Mater. Sci. Forum* **969** 253–9
[2] Gleason K K 2015 Overview of chemically vapor deposited (CVD) polymers *CVD Polymers* (New York: Wiley) pp 1–11
[3] Liu Y, Chen L F, Tang H P, Liu C T, Liu B and Huang B Y 2006 Design of powder metallurgy titanium alloys and composites *Mater. Sci. Eng. A* **418** 25–35
[4] Gui M and Kang S B 2000 6061Al/Al-SiCp bi-layer composites produced by plasma-spraying process *Mater. Lett.* **46** 296–302
[5] Pantelis D, Tissandier A, Manolatos P and Ponthiaux P 1995 Formation of wear resistant Al–SiC surface composite by laser melt–particle injection process *Mater. Sci. Technol.* **11** 299–303
[6] Butola R, Tyagi L, Kem L, Singari R M and Murtaza Q 2020 *Mechanical and Wear Properties of Aluminium Alloy Composites: A Review* (Singapore: Springer) p. 369–91
[7] Chaudhary A, Kumar Dev A, Goel A, Butola R and Ranganath M S 2018 The mechanical properties of different alloys in friction stir processing: a review *Mater. Today Proc.* **5** 5553–62
[8] Butola R, Singari R M, Bandhu A and Walia R S 2017 Characteristics and properties of different reinforcements in hybrid aluminium composites: a review *Int. J. Adv. Prod. Ind. Eng. IJAPIE-SI-MM* **511** 71–80
[9] Yuvaraj N, Aravindan S and Vipin 2015 Fabrication of Al5083/B4C surface composite by friction stir processing and its tribological characterization *J. Mater. Res. Technol.* **4** 398–410
[10] Butola R, Murtaza Q and Singari R M 2020 Formation of self-assembled monolayer and characterization of AA7075-T6/B 4 c nano-ceramic surface composite using friction stir processing *Surf. Topogr.: Metrol. Prop.* **8** 025030
[11] Wang T, Guo W, Wan L and Lv S 2014 Microstructure and surface mechanical property of AZ31 Mg/SiCp surface composite fabricated by direct friction stir processing *Mater. Des.* **59** 274–8
[12] Ramezani N M, Davoodi B, Aberoumand M and Hajideh M R 2019 Assessment of tool wear and mechanical properties of Al 7075 nanocomposite in friction stir processing (FSP) *J. Brazilian Soc. Mech. Sci. Eng.* **41** 182
[13] Bauri R, Ram G D J, Yadav D and Kumar C N S 2015 Effect of process parameters and tool geometry on fabrication of ni particles reinforced 5083 Al composite by friction stir processing *Mater. Today Proc.* **2** 3203–11
[14] Węglowski M S, Sedek P and Hamilton C 2015 Experimental and numerical analysis of residual stress in cast aluminum alloy after FSP process *Key Eng. Mater.* **651–653** 1563–8
[15] Mehdi H and Mishra R S 2020 An experimental analysis and optimization of process parameters of AA6061 and AA7075 welded joint by TIG + FSP welding using RSM *Adv. Mater. Process. Technol.* **6** 1–23
[16] Butola R, Choudhary N, Kumar R, Kumar Mouria P, Zubair M and Singari R M 2021 Measurement of residual stress on H13 tool steel during machining for fabrication of FSW/FSP tool pins *Mater. Today Proc.* **43** Part 1, 256–62
[17] Rathee S, Maheshwari S, Noor Siddiquee A, Srivastava M and Sharma S K 2016 Process parameters optimization for enhanced microhardness of AA 6061/ SiC surface composites fabricated via friction stir processing (FSP) *Mater. Today Proc.* **3** 4151–6
[18] Ahmadkhaniha D, Heydarzadeh Sohi M, Zarei-Hanzaki A, Bayazid S M and Saba M 2015 Taguchi optimization of process parameters in friction stir processing of pure Mg *J. Magnes. Alloy.* **3** 168–72
[19] Butola R, Kanwar S, Tyagi L, Singari R M and Tyagi M 2020 Optimizing the machining variables in CNC turning of aluminum based hybrid metal matrix composites *SN Appl. Sci.* **2** 1356
[20] Bector K, Tripathi A, Pandey D, Butola R and Singari R M 2021 A review on the fabrication of surface composites via friction stir processing and its modeling using ANN *Advances in Manufacturing and Industrial Engineering.* ed R M Singari *et al* (Singapore: Lecture Notes in Mechanical Engineering, Springer) p. 1–11
[21] Ahmad S, Gangwar S, Yadav P C and Singh D K 2017 Optimization of process parameters affecting surface roughness in magnetic abrasive finishing process *Mater. Manuf. Process.* **32** 1723–9
[22] Butola R, Singari R and Murtaza Q 2019 Fabrication and optimization of AA7075 matrix surface composites using taguchi technique via friction stir processing (FSP) *Eng. Res. Express* **1** 025015
[23] Butola R, Jain R, Bhangadia P, Bandhu A, Walia R and Murtaza Q 2018 Optimization to the parameters of abrasive flow machining by taguchi method *Mater. Today Proc.* **5** 4720–9
[24] Bauri R, Yadav D and Suhas G 2011 Effect of friction stir processing (FSP) on microstructure and properties of Al–TiC *in situ* composite *Mater. Sci. Eng. A* **528** 4732–9
[25] Elangovan K and Balasubramanian V 2007 Influences of pin profile and rotational speed of the tool on the formation of friction stir processing zone in AA2219 aluminium alloy *Mater. Sci. Eng. A* **459** 7–18
[26] Elangovan K, Balasubramanian V and Valliappan M 2008 Influences of tool pin profile and axial force on the formation of friction stir processing zone in AA6061 aluminium alloy *Int. J. Adv. Manuf. Technol.* **38** 285–95
[27] Asadi P, Faraji G and Besharati M K 2010 Producing of AZ91/SiC composite by friction stir processing (FSP) *Int. J. Adv. Manuf. Technol.* **51** 247–60
[28] Butola R, Singari R M and Murtaza Q 2020 Mechanical and wear behaviour of friction stir processed surface composite through self-assembled monolayer technique *Surf. Topogr.: Metrol. Prop.* **8** 045007
[29] Mishra R S, Ma Z Y and Charit I 2003 Friction stir processing: a novel technique for fabrication of surface composite *Mater. Sci. Eng. A* **341** 307–10
[30] Paidar M, Asgari A, Ojo O O and Saberi A 2018 Mechanical properties and wear behavior of AA5182/WC nanocomposite fabricated by friction stir welding at different tool traverse speeds *J. Mater. Eng. Perform.* **27** 1714–24
[31] Segal V 2018 Review: modes and processes of severe plastic deformation (SPD) *Materials (Basel).* **11** 1175
[32] Tyagi L, Butola R and Jha A K 2020 Mechanical and tribological properties of AA7075-T6 metal matrix composite reinforced with ceramic particles and aloevera ash via Friction stir processing *Mater. Res. Express* **7** 066526
[33] Ahmad S, Singari R M and Mishra R S 2020 Modelling and optimisation of magnetic abrasive finishing process based on a non-orthogonal array with ANN-GA approach *Trans. Inst. Met. Finish.* **98** 186–98
[34] Ao S, Li C, Huang Y and Luo Z 2020 Determination of residual stress in resistance spot-welded joint by a novel x-ray diffraction *Measurement* **161** 107892
[35] Reghu V R, Ghosh S, Prabhu A, Shankar V and Ramaswamy P 2020 Residual stresses analysis on thermal barrier coatings—NDT tool for condition Assessment *Lecture Notes in Mechanical Engineering. Springer, Singapore* **00** 175–83
[36] Kumar S, Grover S and Walia R S 2018 Effect of hybrid wire EDM conditions on generation of residual stresses in machining of HCHCr D2 tool steel under ultrasonic vibration *Int. J. Interact. Des. Manuf.* **12** 1119–37

Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021)
IEEE Xplore Part Number: CFP21K74-ART; ISBN: 978-0-7381-1327-2

Fake Currency Detection with Machine Learning Algorithm and Image Processing

Aman Bhatia
Department of Software Engineering
Delhi Technological University
Delhi,India
aman.b25@hotmail.com

Vansh Kedia
Department of CSE
Delhi Technological University
Delhi,India
VanshKedia1999@gmail.com

Anshul Shroff
Department of CSE
Delhi Technological University
Delhi,India
anshul.shroff223@gmail.com

Mayand Kumar
Department of Software Engineering
Delhi Technological University
Delhi,India
mayandkumar10@gmail.com

Bickey Kumar Shah
Department of CSE
Delhi Technological University
Delhi,India
Bickeyshah721@gmail.com

Aryan
Department of Computer Science
The University of Sheffield
Sheffield, South Yorkshire, UK
aryan.3@outlook.com

2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) | 978-1-6654-1272-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICCS51141.2021.9432274

***Abstract*—This paper deals with the matter of identifying the currency that if the given sample of currency is fake. Different traditional strategies and methods are available for fake currency identification based on the colors, width, and serial numbers mentioned. In the advanced age of Computer science and high computational methods, various machine learning algorithms are proposed by image processing that gives 99.9% accuracy for the fake identity of the currency. Detection and recognition methods over the algorithms include entities like color, shape, paper width, image filtering on the note. This paper proposes a method for fake currency recognition using K-Nearest Neighbours followed by image processing. KNN has a high accuracy for small data sets making it desirable to be used for the computer vision task. In this, the banknote authentication dataset has been created with the high computational and mathematical strategies, which give the correct data and information regarding the entities and features related to the currency. Data processing and data Extraction is performed by implementing machine learning algorithms and image processing to acquire the final result and accuracy.**

***Keywords*—Accuracy, data extraction, features extraction, image processing, K-Nearest.**

I. INTRODUCTION

In this century where the majority of people are aware of technology and how it works, many of them indulge in unlawful activities. One of such activities is the production of fake currency which is practiced to deceive people. In this proposal, it is focused on this illegitimate practice and try to bring forward a solution for it. According to a survey, the maximum number of cases of counterfeit in India still relate to fake currency, There were 132 cases of counterfeit currency in 2018, which shot up 37 percent to 181 in 2019[7]. In Order to stop this fraudulent activity, a system is proposed that can be integrated into electronic devices that will detect the fake note as soon as it is scanned by the device. Some of the techniques which are considered are used previously and include KNN which will be utilized in the proposed system with enhanced accuracy. K-nearest neighbors (KNN) is an algorithm that stores all the available data and classifies a new data point based on the similarity. This means when new data appears then it can be easily classified into a good suite category by using the KNN algorithm [6]. Usually, the Euclidean distance is used as the distance metric. Then, it assigns the point to the class among its *k* nearest neighbors (where k is an integer).

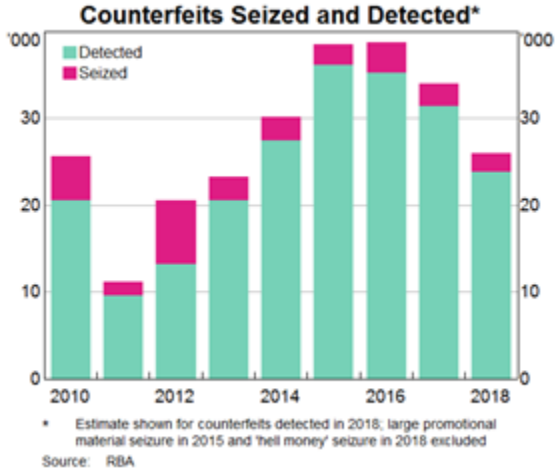


Fig.1 Counterfeits Detected

As *k*-NN does not require the off-line training stage, its main computation is the on-line 'searching' for the *k* nearest neighbors of a given testing example. Although using different *k* values are likely to produce different classification results, 1-NN is usually used as a benchmark for the other classifiers since it can provide reasonable classification performances in many pattern classification problems.
The function used for calculating the Euclidean distance is:

$$dist\,(A,B) = \sqrt{\frac{\sum_{i=1}^{m}(x_i - y_i)^2}{m}}$$

Fig.2 Euclidean formula

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:40:51 UTC from IEEE Xplore. Restrictions apply.

In this straight strategy, some errors and drawbacks were observed that was not making the system and fake identification method more efficient.

- motion blur problem,
- Noise imposed by image capture instrument
- Less efficient feature extraction technique

Due to such problems faced in the major cases the concept for the implementation of image processing works, which purify the entities like shape, color, serial numbers in the form of images and brings more distinction and efficiency in the implementation and work was introduced [3]. As the data set is small this KNN algorithm suits best for the system and high-performance measures scores are expected for the same. Fake currency detection system not only reduces the circulation of counterfeit currency but also provides encouragement to the dealers to accept/make payment through cash, refrains people from circulating fake notes and also ensures proper flow of currency in economy.

II. LITERATURE SURVEY

Different types of study and research work have been carried out in earlier days a different time. Different enhancements and progress were observed [1]. In the past studies the data collected for the fake note detection was with professional cameras but in those data, accuracy seen was to be fair and good due to simple machine learning algorithms. K nearest neighbor algorithms were used traditionally for the detection of fake notes. Systems were getting slower when the data size became large. After that system came across to classify the precision and recognition rate with some enhancement in Machine learning algorithms and deep learning concepts [12]. Due to high and large data sets, data sets were getting distorted, and the precision was not effective a lot though it was 98%. All of these detections were carried out earlier only with open cv and python but time and again with modern deep learning techniques data were collected with the count of 100 images per denomination and then measured [11]. Accuracy of training and testing sets were measured. This brings the chain type efficiency that elongates to a larger value in comparison to other techniques. Concept of the transfer learning was used in the system. The noise was also captured, and this was another problem due to which much more advancement was required. After that, a Convolutional neural network came into the measurement for the error elimination. Loss trends were generally analyzed concerning training loss (TL) and validation loss (VL). Accuracy trends were generally analyzed by training accuracy (TA). In 2021 the fake note is being detected with the algorithms of efficient Machine learning, Deep convolutional neural network, and followed by image processing. It has shown the efficiency to be maximum in today's days.

III. METHODOLOGY

The data set was made by collecting high-quality images of both genuine and counterfeit currency using an industrial camera. The images have a dimension of 400x400 pixels[2]. Due to the type of lens used and the distance to the investigated object, grayscale pictures of 660dpi were captured. On these pictures, Wavelet Transform was used to extract features from the gathered images. The attributes gathered after the Wavelet Transformation were as follows-

- Variance
- Skewness
- Kurtosis
- Entropy
- Class of the currency

First, four of the five derived attributes are continuous and describe the features of the note where-as the fifth attribute- 'Class', classifies the currency into fake by giving it a value of 0 and genuine by giving it a value of 1. The data set has a total of 1372 samples, out of which 610 are genuine and 762 are counterfeit note samples[9].

A brief description of the data set is presented below-

	variance	skewness	kurtosis	entropy	class
count	1372.000000	1372.000000	1372.000000	1372.000000	1372.000000
mean	0.433735	1.922353	1.397627	-1.191657	0.444606
std	2.842763	5.869047	4.310030	2.101013	0.497103
min	-7.042100	-13.773100	-5.286100	-8.548200	0.000000
25%	-1.773000	-1.708200	-1.574975	-2.413450	0.000000
50%	0.496180	2.319650	0.616630	-0.586650	0.000000
75%	2.821475	6.814625	3.179250	0.394810	1.000000
max	6.824800	12.951600	17.927400	2.449500	1.000000

Fig.3 Dataset Description

On analysing the data, following observations was found-

- The data is highly varied. This can be seen as the maximum and mean values of all the features vary greatly.
- Therefore, the data needs to be normalized to prevent from the results being dominated by a single attribute.
- The 'Class' feature has a binary value, 1 denoting a genuine sample and 0 denoting a counterfeit sample.

Furthermore, after plotting each attriubte following can be observed-

- Entropy is negatively skewed which implies that the data has high entropy values[13].
- It is observed that Kurtosis is positively skewed.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:40:51 UTC from IEEE Xplore. Restrictions apply.

- The attributes variance and skewness are evenly distributed across the spectrum.

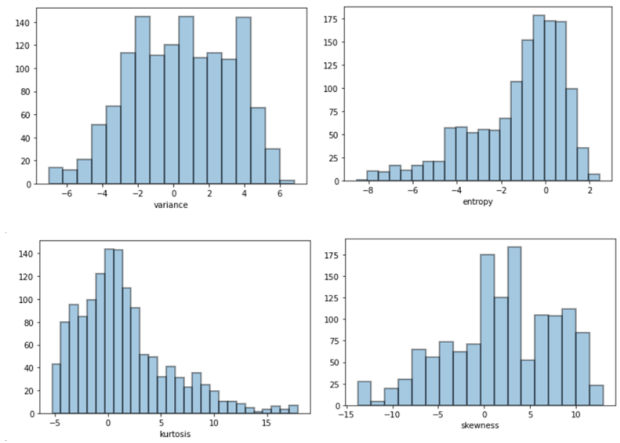


Fig.4 scaling dataset

After glancing at the distribution of the dataset, the apparent observation is that the data needs to be scaled as the values for all the attributes have relatively high variation[14]. This will prevent the model from being biased to a single attribute.

Lastly, by plotting the scatter plots for the given attributes in the dataset the following are observed-

- It can be seen that genuine notes tend to have a low variance value, when compared to the counterfeit notes[8].
- It can be seen from the scatter plot of entropy and kurtosis that, when both combined they perform a very weak task of separating the data. Therefore, any alogrithm used on them will provide with poor results.
- On the other hand, Skewness and variance perform a very good job of segragation the notes into fake and genuine.

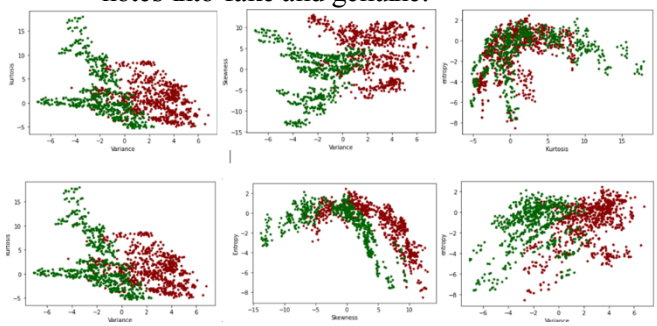


Fig.5 features comparison

Look at the overall dataset, it can be seen that when all the attributes are paired together, they provide good training criteria that help in distinguishing the currency notes[15]. This will help in training a model that can provide higher accuracy. The key highlights from exploring the data set were-

- The data needs to be normalized.
- There are no missing values.
- Outlier detection is not needed to be performed.

A) Normalizing the dataset

After exploration of the data set, It is observed that the data needs to be scaled so that the model is not biassed towards a particular feature. For this, it has been decided to keep the range of each feature between 0 and 1. To do so, the following formula is used-

$$X_{norm} = \frac{X_{current} - X_{min}}{X_{max} - X_{min}}$$

Fig.6 Normalization

To do so, MinMaxScaler is imported from the sklearn. preprocessing library and performed the fit and transform function to the features of the data set. After performing the normalization the data set is looked at below.

	variance	skewness	kurtosis	entropy
0	0.769004	0.839643	0.106783	0.736628
1	0.835659	0.820982	0.121804	0.644326
2	0.786629	0.416648	0.310608	0.786951
3	0.757105	0.871699	0.054921	0.450440
4	0.531578	0.348662	0.424662	0.687362
5	0.822859	0.877275	0.057100	0.489711

Fig.7 Normalization of dataset

B) Algorithms Used to train the model-

Training of the data can be carried out with the different functionalities and algorithm. For the purpose of this project,it is decided to compare the working of 3 algorithms and perform a comparative study on them.

- K- Nearest Neighbours
- Support Vector Classifier
- Gradient Boosting Classifier

1) K-Neartest Neighbors- It is a lazy learning supervised algorithm. The reason it is called lazy is because it doesn't have a specialized training phase and uses all the data for training while classification. It is mostly used for classification predictive problems. The way the algorithm works is that it classifies a given data point by looking at its closest neighbour and assigns a weight to them based on the distance. The distance can be Euclidian, Manhattan,etc, and it gives more weightage to the closer data points. There are a couple of advantages which include- it is simple to implement, flexible to the types of features and naturally handles multi classes. However, the drawback is that a lot of heavy computation work happens while testing which makes it a slow algorithm.

For the data set used, KNN is a good option as the data set isn't huge so the computation time is not a major

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:40:51 UTC from IEEE Xplore. Restrictions apply.

issue. Moverver, as it is non-parametric and do not need to worry about the data distribution and it can the decision boundary can take any form[16].

2) Support Vector Classifier- It is a supervised machine learning algorithm which is mostly used for classficiation and regression problems. In this algorithm Each point is plot in a n-dimensional plane with the value of the point being the data point of the sample. Then the classification is performed by finding a hyper-plane that is the best fit for dividing the classification. Some advantages of this algorithm are- performs well for non linear decision boundaries, uses a subset of training points so is memory efficient, and is versatile. Some of the drawbacks include- if the number of features are greater than the number of sample the algorithm is likely to underperform and it is quite taxing to train[5].

This algorithm can work on our data set as it is classified into 2 classes and the number of features is quite less that the number of samples. Moreover, as discussed above, even 2 features provide us with some clear distinction between counterfeit and genuine notes. Therefore, after projecting to a higher dimensional plane we should get better results.

3) Gradient Boosting Classifier- It is an ensemble of weak learner which when work together help in providing a very accurate predicting model. A weak learner are algorithms which provide models with accuracy being slightly better than randomly choosing an answer. Therefore, they are structured using a decision tree type model in which layers of yes and no type questions are asked to get a predictive model. Some of its advantages are that it is easy to interpret and implement, can effectively capture non-linear function dependencies, it is extremely flexible and cutomizable and finally builds a robust model using the output of many weak learners. However, they consume a lot of memory and can be very time consuming with the ensemble is large.

This algorithm can be used on our data set as layered questions can be asked and the data and be classified according. Moreover, as the number of features are less the training and testing time is not going to huge.

C) Implementation

For this project the steps we are going to follow are as follows-

- Preprocess and normalize the data set.
- Split the data using K-fold cross validation technique.
- Training the model using the algorithms discussed above.
- Calculate the performance measures.

Initially, the data set was loaded and preprocessed. As discussed in the above sections data normalization was required that was done using the MinMaxScaler present in the sklearn. preprocessing library. Independent and Dependent Variables- As discusses above the data set contains 5 columns with 4 being the features extracted after Waveform transform and the 5th one being the class that distinguishes between fake and genuine notes. For this project, we took the first 4 attributes-'variance', 'skewness', 'kurtosis' and 'entropy' as the independent or input variables and the 'class' as the target or dependent variable. Data is split after the normalization and preprocessing works. We used the Kfold functionality present in the sklearn.model_selection to split the data. The purpose of using this library is it helps to gain a more accurate understanding of the algorithm. As the data is split into k blocks of data set and the algorithm is run with each block once as the test set. This helps in gaining a complete understanding of the model as each sample is used for both training and testing. Another reason we used this method was that our data set was limited and therefore, didn't take too much extra time in training through this method.

Once the data was split into training and testing sets, we moved on to building the model.

D) Building the Model

For this project, the required classifiers were imported from the sklearn library. We made a function that took parameters as the training and testing data sets along with the algorithm and returned a trained model with its performance measure scores. We repeated this for every fold and accumulated results for every pass in a list. Lastly, the cross-validation predict module is used to build a confusion matrix for all the algorithms. Finally, plotted the results to gain a visual understanding of how the algorithms performed[4].

E) Performance measures

Performance measures are used to check the correctness of the model. For this project, accuracy, precision, and f-score are performance measures. graphs were plotted for each performance measure and compared the results. k-fold cross-validation technique is used for this project. As discussed above the KFold and cross-validation predict module from the sklearn model selection library. This technique divides the data set into the number of 'k' blocks specified by the user and runs the algorithm once for every block as a test set 10 for the value of k. After building the model it is analyzed using the performance measures discussed above.

IV. RESULT ANALYSIS

After training the model using the data set and the algorithms described above, we tested them to see how they performed. Calculating the performance measures described above, we were able to gain an understanding of how every algorithm performed. Below are the accuracy, precision, and f-scores for each algorithm.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:40:51 UTC from IEEE Xplore. Restrictions apply.

Algorithm	Accuracy	Precision	F-Score
KNN	99.9%	99.9%	99.9%
SVC	97.5%	99.7%	98.6%
GBC	99.4%	99.9%	99.7%

Table.1 Algorithm Comparison accuracy

Below is the graphical representation of the above results,

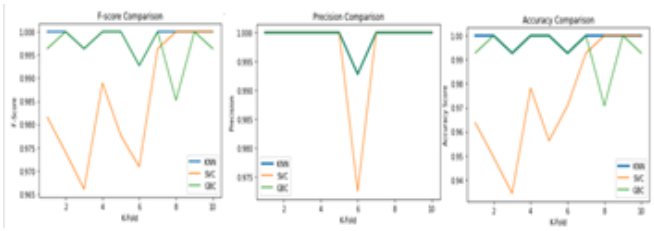


Fig.8 Algorithm Comparison

KNN			SVC			GBC	
760	2		736	26		759	3
0	610		0	610		3	607

Table.2 Matrix

As can be seen from the above results, KNN performs with the most consistency. It has the lowest accuracy of 99.2%., however, 80% of the time it gave a result with 100% accuracy. Comparing the other two algorithms, it can be seen that GBC was the closest to the performance of KNN with an accuracy of 99.4% while SVC had the lowest result of 97.5%. However, it can be noted that all three algorithms had an accuracy above 97% which is quite impressive. Similar results can be derived by comparing the results for precision and f-score. Moreover, taking a look at the confusion matrices, it can be noticed that KNN gave only 2 wrong predictions whereas GBC gave 6 and SVC predicted 26 samples incorrectly. This consolidates the fact that KNN outperforms the rest of the algorithms for this case. As such a project needs to have high accuracy as predicting even some notes as false positives or negatives can cause major faults, it will not be possible to use the model build by using SVC as it produces 26 faulty predictions. Looking at these results it can be seen that for the given data set the most accurate algorithm to use is KNN. Moreover, it can be noted that KNN predicted all genuine notes correctly, which is essential in the real world as predicting counterfeit as an authentic currency will be more detrimental.

V. CONCLUSION AND FUTURE WORK

After implementing and analyzing the results gathered, we can deduce that all the three algorithms used were exceptionally accurate at classifying notes as genuine and counterfeit based on the used data set. However, KNN outperformed the other two as discussed in the above section. It had an accuracy of 99.9% with classifying incorrectly classifying only 2 counterfeit notes. However, this result is limited as the data set used was quite small. It had a total of 1372 samples which when considered in the real-world scenario might not perform as well as it has currently. To build on this, we propose to form a much larger data set with real-world like pictures of real and fake currency notes. This will help in providing a much more realistic model. Moreover, with a large data set available, deep learning algorithms like Convolutional Neural Networks or CNN can be applied which have high accuracy in image processing scenarios. Furthermore, by using CNN the project can directly analyze images as input, and wavelet transformation will not be required. This can make the system more convenient and user friendly to use. Moreover, as it is likely to be used in financial institutions it will be more convenient for users to directly click a picture and get it verified, this can be done with the help of CNN as mentioned above. Hence, to make the project more robust and professional the above suggest measures can be implemented[10].

ACKNOWLEDGMENT

The Authors are indebted to the central library of Delhi Technological University, Delhi India, for providing all the access to research materials with the lab required to successfully complete the research work.

REFERENCES

[1] D. V. Kapare, S. Lokhande, and S. Kale, “Automatic Cash Deposite Machine With Currency Detection Using Fluorescent And UV Light,” vol. 3, pp. 309–311, 2013.

[2] M. N. Rathore and J. Sagar, “A Review on Fake currency detection using feature extraction,” vol. 10, no. 11, pp. 407–411, 2019.

[3] K. Santhanam, S. Sekaran, S. Vaikundam, and A. M. Kumarasamy, “Counterfeit currency detection technique using image processing, polarization principle and holographic technique,” *Proc. Int. Conf. Comput. Intell. Model. Simul.*, no. Figure 2, pp. 231–235, 2013, doi: 10.1109/CIMSim.2013.44.

[4] P. Ponishjino, K. Antony, S. Kumar, and S. Jebakumar, “Bogus currency authorization using HSV techniques,” *Proc. Int. Conf. Electron. Commun. Aerosp. Technol. ICECA 2017*, vol. 2017-January, pp. 179–183, 2017, doi: 10.1109/ICECA.2017.8203667.

[5] Q. Zhang and W. Q. Yan, “Currency Detection and Recognition Based on Deep Learning,” *Proc. AVSS 2018 - 2018 15th IEEE Int. Conf. Adv. Video Signal-Based Surveill.*, pp. 0–5, 2019, doi: 10.1109/AVSS.2018.8639124.

[6] A. Upadhyaya, V. Shokeen, and G. Srivastava, “Analysis of counterfeit currency detection

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:40:51 UTC from IEEE Xplore. Restrictions apply.

techniques for classification model," *2018 4th Int. Conf. Comput. Commun. Autom. ICCCA 2018*, pp. 1–6, 2018, doi: 10.1109/CCAA.2018.8777704.

[7] S. Arya and M. Sasikumar, "Fake Currency Detection," *2019 Int. Conf. Recent Adv. Energy-Efficient Comput. Commun. ICRAECC 2019*, pp. 2019–2022, 2019, doi: 10.1109/ICRAECC43874.2019.8994968.

[8] A. Kumar and A. Kumar, "Dog Breed Classifier for Facial Recognition using Convolutional Neural Networks," pp. 508–513, 2020.

[9] M. Haider Ali, "Thesis Report on Fake Currency Detection using Image Processing Method," *Akiful Mohaimin Rifat Islam Shahriar Chowdhury*, no. 13301148, pp. 1–38, 1330.

[10] M. N. Shende and P. P. Patil, "A Review on Fake Currency Detection using Image Processing," *Int. J. Futur. Revolut. Comput. Sci. Commun. Eng.*, vol. 4, no. 1, pp. 391–393, 2018.

[11] T. Agasti, G. Burand, P. Wade, and P. Chitra, "Fake currency detection using image processing," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 263, no. 5, pp. 88–93, 2017, doi: 10.1088/1757-899X/263/5/052047.

[12] P. . P. Binod Prasad Yadav, C. S. Patil, R. R. Karhe, "An automatic recognition of fake Indian paper currency note using MATLAB," *Certif. Int. J. Eng. Sci. Innov. Technol.*, vol. 9001, no. 4, pp. 2319–5967, 2008, [Online]. Available: http://www.ijesit.com/Volume 3/Issue 4/IJESIT201404_77.pdf.

[13] M. A. Gaikwad, V. V Bhosle, and V. D. Patil, "Automatic Indian New Fake Currency Detection Technique," *Int. J. Eng. Res. Technol.*, vol. 6, no. 11, pp. 84–87, 2017.

[14] S. Adhikari, S. Thapa, and B. K. Shah, "Oversampling based Classifiers for Categorization of Radar Returns from the Ionosphere," *Proc. Int. Conf. Electron. Sustain. Commun. Syst. ICESC 2020*, no. Icesc, pp. 975–978, 2020, doi: 10.1109/ICESC48915.2020.9155833.

[15] A. Ghimire, S. Thapa, A. K. Jha, S. Adhikari, and A. Kumar, "Accelerating business growth with big data and artificial intelligence," *Proc. 4th Int. Conf. IoT Soc. Mobile, Anal. Cloud, ISMAC 2020*, pp. 441–448, 2020, doi: 10.1109/I-SMAC49090.2020.9243318.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:40:51 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021)
IEEE Xplore Part Number: CFP21K74-ART; ISBN: 978-0-7381-1327-2

Feature Extraction Techniques in Facial Expression Recognition

1st Nitendra Mishra
Department of Computer Science Engineering,
Delhi Technological University
Delhi, India
nitendramishra$_2$k19cse16@dtu.ac.in

2nd Aruna Bhatt
Department of Computer Science Engineering.)
Delhi Technological University.)
Delhi, India
aruna.bhat@dtu.ac.in

2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) | 978-1-6654-1272-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICCS51141.2021.9432192

***Abstract*—Facial expression recognition became a very important and very challenging aera within past few years in computer science vision field. It (FER) a very hot and active topic for research. FER play interesting role in aera of computer vision processing.This paper reviews the different algorithms which is used in facial expression recognition. usually, Facial expression recognition has 3 Majour stages first one is preprocessing I.e., face detection second is Feature extraction or selection (the aim of this paper) and last one is classification. In this paper study we review some current work done for Facial expression recognition technique. we know that feature extraction method is very important step for image recognition so in this paper we see that after the preprocessing the images how feature extraction done, we see some technique like local binary pattern (LBP), Gabor filter etc. And also, this paper gives an evaluation of Majour characteristic feature extraction and based on that their performance in terms of accuracy and them limitations.arrangement of this FER paper is following in order: section 1 contain the introduction section II contains a sample process of facial expression recognition section III contains the different technique of feature extraction and section IV contains the review discussion and last section V contain the conclusion.to making these algorithms efficient and more effective feature extraction technique played an important role.**

***Index Terms*—FER, Face Expression Recognition, Face Emotion Recognition, Computer Vision, Facial Expression Recognition, Feature Extraction**

I. Introduction

Facial expression recognition software system may be a technology that uses biometric markers to discover emotions in human faces. additional exactly, facial expression technology may be a sentiment analysis tool and is ready to mechanically discover the six basic expressions which is shown in fig 1 I.e., anger, disgust, fear, happiness, sadness and surprise. Facial expression recognition in present time widely used in various research field such as eLearning, and one of the most important factors in traffic accidents i.e., driving fatigue we used facial expression recognition for driving fatigue detection and many human computer interfaces and so many other research fields[1] to [5].so basically, we developed some



Fig. 1. six basic expressions I.e., anger, disgust, fear, happiness, sadness and surprise

algorithm for detecting our six major's expression and it's a very interesting and hot topic. There are three main steps for facial expression recognition the first one is prepressing[6] it's a very formal step it has many work one of them is removal all noise from image done here and we resize the images here. Now second one is feature extraction (we discuss all in this paper) basically its extract all major information of images here and last is classification[7] which is again most important step for facial expression recognition. Classification basically groups the images on basic of their features they belong.in previous some year we have improved some technique in face detection and also done great improvement work in feature extraction and classification [8] to [12]. we found that if we

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:41:38 UTC from IEEE Xplore. Restrictions apply.

want to increase our recognition rate then we use fuzzy logic [13], neural networks [14], deep neural networks[15] etc. with facial expression recognition algorithms. During apply these algorithms we have to very careful because there are many factors that work as obstacles like changes of position of images, dim lighting, changes of expression etc. Dealing with these kinds of obstacles is really a big challenge.

This paper reviews the different algorithms which is used in facial expression recognition. We can see that there are typically two methods for FER the first one is which includes the all technique of feature based[16] and second one is model based[17].in feature-based model further we can divide it into two sub group i.e., geometric based and appearance based.in geometric based model for represent location of images like for ear nose etc. we use some pointer curves etc. we know that feature extraction method is very important step for image recognition so in this paper we see that after the preprocessing the images how feature extraction done we see some technique like local binary pattern (LBP) , Gabor filter etc. he arrangement of this FER paper is following in order: section 1 contain the introduction section II contains a sample process of facial expression recognition section III contains the different technique of feature extraction and section IV contains the review discussion and last section V contain the conclusion

II. Facial Expression Recognition Process

Facial recognition process has multiple phases as we discuss in section one the first one is preprocessing and second one is facial detection and features extraction and last one is classification; we can see it on fig 2.

A. Preprocessing

If you are dealing with images processing then the first step is always preprocessing. Whatever technique we are using we always use preprocessing as a first step. When we start the image processing, we set a criterion for image input we call it as a standard form for input we required our image always in standard form because it might chance that the input image not fulfilled our criteria may be not fit in size, in color we want might chance that it affected with noise, affected background and illumination variants[18][19] etc.so that is the reason the first step of preprocessing is required for filtering the noise removing the extra thing which is not required in image also fit the size of the images so that it can become a standard form for input etc.

B. Detection of face

When the first step i.e., preprocessing is done then our image (standardized image) pass through the detection of face phase and here in this phase we remove the unusual background and try to extract the solo image face. For doing this operation some techniques help which are edge detection techniques, Sobel techniques, canny , laplacian and many others

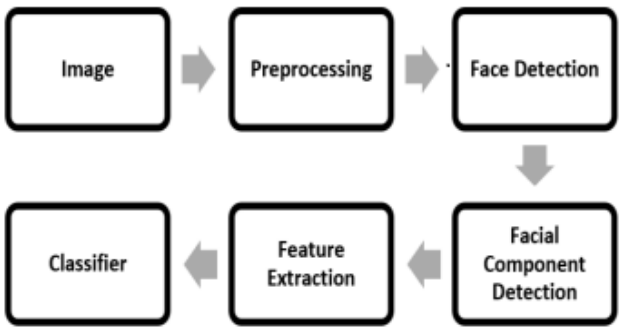


Fig. 2. General Step for Facial Expression Recognition

C. FCR (Facial Component Revealing)

feature extraction method is very important step for image recognition. The efficiency of FER is based on this phase how it works. The technique which are used in feature extraction are local binary pattern, Gabor filter etc. We'll see all features extraction details.

D. feature extraction

eature extraction method is very important step for image recognition. The efficiency of FER is based on this phase how it works. The technique which are used in feature extraction are local binary pattern, Gabor filter etc. We'll see all features extraction details.

E. Classifier

After pass the all above layer classifier technique classifies the image to its appropriate class level expression. There are some classifications techniques are Deep Neural Networks, support vector machines, Neural Networks, etc.[20].

III. FEATURE EXTRACTION METHODS

As we earlier discuss about feature extraction now there are some techniques of feature extraction that discuss some of feature extraction technique are histograms of oriented gradient, geometric features, local Binary Pattern (LBP), Gabor filter texture, learned feature and hybrid feature.

A. Histogram of Oriented Gradient

HOG may be a descripter of feature which is used in many field where object shape is matters. When need of object detection in field of computer vision then histogram of oriented gradient is playing a very important role.basically its count the number of occurrence of gradient orientedin localizede portioon of an image.in histogram of oriented gradient try to find out the a bar chart of gradient direction for doing this input(image) is devided into the tiny connection parts which refers to the cell and pixel inside the every cells. Sincee totally pictures might have different contrast thus, distinction standardization is critical to boost performance.this standardization leads to higher unchangeability to change in illumination.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:41:38 UTC from IEEE Xplore. Restrictions apply.

B. Geometric Features

Geometric features are basically used for solving visual related problems. Its(geometric) the combo of both machine learning and computer vision. The primary focous of this technique is to search out a group of representtive options of geommetric kind to repreesent Associate in Nursing object by grouping geometric options from pictures and learning them mistreatment economical machine learrning strategies. It(geometric) options don't seems to be affecteed by lightingg condition. there are no any troublesome to perform any action. Ghimire and le[21] they extract some method of geometric features mechanically from video frame victimization displacements supported by elasticity bunchh graph matchhing displacment estimaation.

C. Local Binary Pattern

Local Binary Pattern (LBP) may be a straightforward nonetheless terribly economical texture operator that label the components of a picture by threholding the neighborhood of every pixels and consider the result as a binary range.local binary pattern uses (3x3) pixel that contains grey scale values and threshhold each Neighbour element P(0 to7) with the middle element R(1) to come up with a binarry sequence employing a binary threholding perform given in 1 and then cypher the deciimal equivalentt for the middle element with 2. figure 3 show is procedure: Suppose general practitioner is that the gray values in constituent P(0 to 7) and rate is the gray value of the middle constituent R(1) so, the function of edge is:

$$s(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases} where \;\; x = g_p \times g_c \tag{1}$$

$$LBP_{R,P} = \sum_{p=0}^{p-1} s(x)2^p \tag{2}$$

Based on the image higher than, we are able to extract the bar chart of every region as follows: As we've got a picture in gray-scale, every bar chart (from every grid) can contain solely 0 to 255 = 256 position represent the occurrencces of every picture element intensity. so, we want to concatenate every bar chart to form a brand new and larger bar chart. Suppose we've got 8x8 grid, we'll have (8x8x256)=(16.38) position within the final bar chart. the ultimate bar chart represents the characteristis of the original image.like this[22] administrated a small research study on local binary pattern with linear programming options victiization (JAFFE) information. According to study it shows the results that median accuracy of ninety-three.8percentage was achieved.

LBP really have challenges with some factors like; rotation, will increase in procedure quality because the size of options will increase in term of your TIme and house, lower sampled size and restricted info illustration, cause it doesn't take into account magnitude info however component difference. These de-merit Leads to its variants, [23] used one of Local binary pattern variant referred to as (CLBP) that thought of the sign and magntude info of the variations between the middle and therefore the neighboaur grey values.

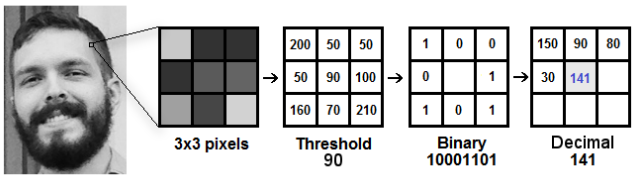


Fig. 3. General Step for Facial Expression Recognition

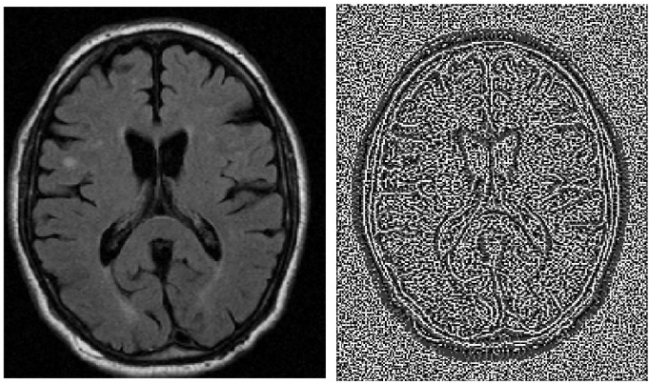


Fig. 4. input image and the corresponding LBP image

D. Gabor Filter Texture

In computer vision field Gabor is very popular for extracting the features. GFT uses the linear filter which is commonly used for texture analysis. The main work of it is to analyzes that any particular freq. Contain in the data (input image) in specific direction in a localized part around the region of analysis. Gabor filter is generally outlined in equ. 3 and 4

$$G(x,y) = \frac{1}{2\pi s_x s_y} exp\left[-1/2\left(\frac{x'^2}{s_x^2} + \frac{y'^2}{s_y^2}\right)\right] exp\left[j\frac{2\pi\bar{x}}{\lambda}\right] \tag{3}$$

$$Where \;\; \bar{x} = x\cos\theta + y\sin\theta, \bar{y} = -x\sin\theta + y\cos\theta \tag{4}$$

Here in equation 3 and 4 forward subsequent parameters: (x y) would be the position of element within any special kind of domain ie . Here the symbol is the wavalenghth in pixel. And also we can see in equation that we have is the orientaion of the it that is our gabor filter and also sx and sy is S.D(standard deviation) in the x and y direction.

E. Hybrid Features

Hybrid options provide space to the analysis some ques. of however great will options be merge for aftmost performance. According to Liu et al.[24] they used the feature extraction mixture of local binary pattern and histogram of oriented gradient from (CK+) and (JAFFE) information and its reduce the extracte options spatial property with (PCA) aftear permutation the mixture on many classifeer he find that the mixed options on soft-max classifier created ninty eight.3percentage on (ck+) and nintieth on (JAFFE) information. So for this we

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:41:38 UTC from IEEE Xplore. Restrictions apply.

can conclude that if we use a perfect(proper) hybrid options that can improve our system performance.

TABLE I
TABLE FOR ACCURACY WITH DIFFERENT DATA-SET AND FEATURE

Author Data set	FET(Features)	Accuracy
Feng Et Al JAFFE(POSED)	LBP	93.8
D huang CK+(POSED)	LBP	84.6
Y Liu JAFFE(POSED)	Hybrid(LBP+HOG)	98.3
K.Sikka OULUCASIA VIS	HMM	75.62
G.SandBach BU4DFE(spontaneous)	3D motion based feature	81.93
P kumar CK+(POSED)	HOG	95
D anggraeni POSED JAFFE POSED	hand crafted preprossing stages	97.06

IV. DISCUSSION

From table 1 used of various classifier LBM produces 93.8 percentage the limitation using this is Postured dataset for system preparing will mislead the framework in real time environment, 84.6 percent accuracy with data set JAFFE (posed) and CK+ (posed) limitation with this data set is fortune of relevant information is conceivable . Hybrid (LBPHOG) produces 98.3 percent Investigate is still finding best combinations of features and 90 percent accuracy with data set CK+(posed) JAFFE (posed) limitation is Propensity to overfit reason Restricted data for preparing. . likewise, HMM produces 75.62 percent accuracy with data set OULU-CASIA VIS (Spontaneous) using SVM classifier perform well and when we used hand crafted prepossessing stages produces the accuracy of 97.06 and with data set ck+ in posed and JAFFE in posed and MUG with posed.Both Cross breed highlights and combination of the hand made highlight and significant learning create a promising execution at that point more center ought to be given to finding the best combination of either hand created highlights or combination of both hand made highlight and learned highlights. We can ascertain that when we use the spontneous data set then we record some lower accuracy compare to when we use some posed data set.

V. CONCLUSION

Machine learning AI covers numerous parts of information, like likelihood hypothesis, measurements, and so on, which is a vital examination bearing in man-made consciousness.In spite of the fact that it may see like profound learning methods for include extraction are more vigorous to scale, impediment, misshapening, turn, etc and have pushed the limits of what was conceivable utilizing conventional computer vision strategies doesn't cruel the computer vision strategies are out of date. The use of profound figuring out how to FER framework ease individuals of the pressure from include extraction and arrangement issues, in light of the fact that these are appropriately thought of. All things considered, profound learning is tested with size of data set yet this could be represented utilizing some information designing methods on the accessible data set. Likewise hand made strategies can be utilized as extra data that could direct an organization in forecast.we know that deep learning has its own challenges in terms of its data set size but we can handle this with some technique which is based on data engineering.Future work is centered on moving forward the data sets and diminishing the preparing time for video preparing[25].

REFERENCES

[1] S. H. Abdurrahim, S. A. Samad, and A. B. Haddin, —Review on the effects of age, gender, and race demographics on automatic face recognition, ‖ Vis. Compute., vol. 34, no. 11, pp. 1617–1630, 2018
[2] G. M. O. Zebari, K. Faraj, and S. R. Zebari, —Hand Writing Code-PHP Or Wire shark Ready Application Over Tier Architecture with Windows Servers Operating Systems or Linux Server Operating Systems, ‖ 2016.
[3] O. Alzakholi, L. Haji, H. Shukur, R. Zebari, S. Abas, and M. Sadeeq, —Comparison Among Cloud Technologies and Cloud Performance, ‖ J. Appl. Sci. Technol. Trends, vol. 1, no. 2, pp. 40–47, Apr. 2020, doi: 10.38094/jastt1219.
[4] L. M. Haji, R. R. Zebari, S. R. M. Zeebaree, M. A. WAFAA, H. M. Shukor, and O. Alzakholi, —GPUs Impact on Parallel Shared Memory Systems Performance – International Journal of Psychosocial Rehabilitation,‖ May 22, 2020. https://www.psychosocial.com/article/PR280814/24791/ (accessed May 22, 2020).
[5] O. H. Jader, S. R. Zeebaree, and R. R. Zebari, —A State Of Art Survey For Web Server Performance Measurement And Load Balancing Mechanisms.‖
[6] D. A. Zebari, H. Haron, S. R. Zeebaree, and D. Q. Zeebaree, —Multi-Level of DNA Encryption Technique Based on DNA Arithmetic and Biological Operations,‖ in 2018 International Conference on Advanced Science and Engineering (ICOASE), 2018, pp. 312–317.
[7] A. A. Salih and M. B. Abdulrazaq, —Combining best features selection using three classifiers in intrusion detection system,‖ in 2019 International Conference on Advanced Science and Engineering (ICOASE), 2019, pp. 94–99.
[8] S. Li and W. Deng, —Deep facial expression recognition: A survey,‖ IEEE Trans. Affect. Comput., 2020.
[9] O. F. Mohammad, M. S. M. Rahim, S. R. M. Zeebaree, and F. Y. Ahmed, —A survey and analysis of the image encryption methods,‖ Int. J. Appl. Eng. Res., vol. 12, no. 23, pp. 13265–13280, 2017.
[10] B. S. Osanaiyeet al., —Network Data Analyser and Support Vector Machine for Network Intrusion Detection of Attack Type.‖
[11] R. R. Zebari, S. R. Zeebaree, and K. Jacksi, —Impact Analysis of HTTP and SYN Flood DDoS Attacks on Apache 2 and IIS 10.0 Web Servers,‖ in 2018 International Conference on Advanced Science and Engineering (ICOASE), 2018, pp. 156–161.
[12] R. R. Zebari, S. R. Zeebaree, K. Jacksi, and H. M. Shukur, —E-Business Requirements For Flexibility And Implementation Enterprise System: A Review.‖
[13] R. Zebari, A. Abdulazeez, D. Zeebaree, D. Zebari, and J. Saeed, —A Comprehensive Review of Dimensionality Reduction Techniques for Feature Selection and Feature Extraction,‖ J. Appl. Sci. Technol. Trends, vol. 1, no. 2, pp. 56–70, May 2020, doi: 10.38094/jastt1224
[14] S. R. Zeebaree, K. Jacksi, and R. R. Zebari, —Impact analysis of SYN flood DDoS attack on HAProxy and NLB cluster-based web servers,‖ Indones. J. Electr. Eng. Comput. Sci., vol. 19, no. 1, pp. 510–517, 2020.
[15] U. Scherhag, C. Rathgeb, J. Merkle, R. Breithaupt, and C. Busch, —Face recognition systems under morphing attacks: A survey,‖ IEEE Access, vol. 7, pp. 23012–23026, 2019
[16] D. A. Zebari, H. Haron, S. R. Zeebaree, and D. Q. Zeebaree, —Enhance the Mammogram Images for Both Segmentation and Feature Extraction Using Wavelet Transform,‖ in 2019 International Conference on Advanced Science and Engineering (ICOASE), 2019, pp. 100– 105.
[17] W. M. Abduallah and S. R. M. Zeebaree, —New Data hiding method based on DNA and VigenereAutokey,‖ Acad. J. Nawroz Univ., vol. 6, no. 3, pp. 83–88, 2017

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:41:38 UTC from IEEE Xplore. Restrictions apply.

[18] I. A. Khalifa, S. R. Zeebaree, M. Ataş, and F. M. Khalifa, —Image Steganalysis in Frequency Domain Using Co-Occurrence Matrix and Bpnn,‖ Sci. J. Univ. Zakho, vol. 7, no. 1, pp. 27–32, 2019.
[19] M. A. Mohammed et al., —An anti-spam detection model for emails of multi-natural language,‖ J. Southwest Jiaotong Univ., vol. 54, no. 3, 2019.
[20] A. Fathima and K. Vaidehi, —Review on Facial Expression Recognition System Using Machine Learning Techniques,‖ in Advances in Decision Sciences, Image Processing, Security and Computer Vision, Springer, 2020, pp. 608–618.
[21] D. Ghimire and J. Lee, "Geometric Feature-Based Facial Expression Recognition in Image Sequences Using Multi-Class AdaBoost and Support Vector Machines," Sensor, vol. 13, pp. 7714–7734, 2013.
[22] X. Feng, M. Pietikäinen, and A. Hadid, "Facial Expression Recognition Based on Local Binary Patterns 1," Pattern Recognit. Image Anal., vol. 17, no. 4, pp. 592–598, 2007.
[23] A REVIEW ON FACIAL EXPRESSION RECOGNITION SYSTEMS S. Shaul Hammed1* , Dr.A. Sabanayagam2, E. Ramakalaivani3 ISSN-2394-5125 Vol 7, Issue 4, 2020
[24] F. Ahmed, H. Bari, and E. Hossain, "Person-Independent Facial Expression Recognition Based on Compound Local Binary Pattern (CLBP)," Int. Arab J. Inf. Technol., vol. 11, no. 2, pp. 195–203, 2014
[25] Kumar, T. Senthil. "Video based Traffic Forecasting using Convolution Neural Network Model and Transfer Learning Techniques." Journal of Innovative Image Processing (JIIP) 2, no. 03 (2020): 128-134

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:41:38 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Fifth International Conference on Trends in Electronics and Informatics (ICOEI).
IEEE Xplore Part Number:CFP21J32-ART; ISBN:978-1-6654-1571-2

FPGA Implementation of a pipelined and pseudo-randomized TDES algorithm

Arnab Roy
Dept. of Electronics and Communication Engineering
Delhi Technological University
Delhi,India
arnab6749@gmail.com

Anubhav Gadhok
Dept. of Electronics and Communication Engineering
Delhi Technological University
Delhi,India
anubhav2199@gmail.com

Deva Nand
Dept. of Electronics and Communication Engineering
Delhi Technological University
Delhi,India
devkamboj07@gmail.com

2021 5th International Conference on Trends in Electronics and Informatics (ICOEI) | 978-1-6654-1571-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICOEI51242.2021.9452952

***Abstract*—In today's world where data breaching and misuse of confidential data are so common, an end user's utmost priority should be to keep one's data secure. This is where information security and cryptography comes into the picture and in this paper, a pipelined and pseudo randomized implementation of TDES encryption standard on FPGA using Verilog HDL is designed which can encrypt and decrypt a given user's messages with high throughput and the pseudo randomizer randomizes the key with the help of a Linear Feedback Shift Register which makes the encrypted output random thereby making it more secure than the traditional TDES systems. The target device used is xc7vx330t-3ffg1157 from Virtex-7 FPGA family offered by Xilinx on which a maximum frequency of 608.224 MHz and maximum throughput of 38.92 Gigabits per second was achieved. The proposed design is also faster when compared to previous works on their respective platforms. The proposed system was designed and analyzed thoroughly using the Xilinx ISE Design Suite v14.7.**

***Keywords*—Cryptography,Encryption,TDES,FPGA,Pipelining, Throughput,Information Security**

I. Introduction

With the approach of innovation, the information which is sent and put away in advanced world is expanding step by step. On the other hand, this data is also at a high risk of being accessed unethically. The threat to data security continues to increase exponentially as new types of advanced attacks on data are emerging regularly. To prevent all these mishaps, data needs to be protected so that it can be accessed by only those people who are authorized to do so. This is where encryption comes into picture. Encryption is the process of converting data into some other form which is unrecognizable by unauthorized or normal people and the reverse process is called Decryption [1]. The data is converted into a form which is unreadable by normal people and only the authorized people can convert this unreadable form into readable form. A lot of encryption methodologies are used in the current digital world like DES, TDES, AES, RSA, AES, ECC, Blowfish, Twofish, Threefish, RC5, IDEA [2] [3]. All these algorithms have their own advantages and disadvantages which decide their usage in real time applications.

This paper implements a pseudo randomized pipelined version of TDES in ECB mode with a modified key generation function to make it more secure and efficient than the regular TDES. In order to implement TDES [4], DES algorithm [5] is implemented first followed by introduction of pipelined architecture into the existing algorithm [6] [7] which in return makes the encryption and decryption process faster than before with higher throughput. In order to improve the security, a modified key generation algorithm is employed which consists of sequence of pseudo random numbers generated by Linear Feedback Shift Registers [8] [9] masking the generated key thereby producing a pseudo randomized enciphered output which is completely foreign for an unauthorized user. On the other hand, an authorized user who has access to the master key as well as the seed to the LFSR can safely access the data. These modified DES blocks (both encryption and decryption) are then used to construct a pipelined version of TDES which is faster and more secure than many of the current DES and TDES based systems [10] [11] [12].Cryptography has come a long way over years [13] and in the present world this system can be used in e-governance and banking security [14], smart card applications [15], cloud security [16], etc. The paper organization is as follows: Section II describes the methodology used to design the pipelined DES block. Section III explains the modified key generation employed in this paper which helps in the increasing security of the proposed system. Section IV elucidates the detailed implementation of the final design. Detailed comparisons, relevant results and conclusions are drawn in the final sections of the paper.

II. Pipelined DES Implementation

The Data Encryption Standard (DES) is a symmetric key encryption algorithm which is used for encryption and protection of computerized information. Despite the fact that its key length is short (56 bits) which makes it highly insecure for applications, it has been profoundly influential in the progression of cryptography. In order to implement a pipelined TDES design , it is necessary to construct a coherent pipelined DES algorithm first. The complete design and description of the DES algorithm is presented in [5]. This section will show the pipelining introduced in the existing DES algorithm.

Pipelining increases the throughput of the system by means of parallel computation [7]. Instead of processing intermediate data one by one, pipelining helps in processing intermediate

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:58:51 UTC from IEEE Xplore. Restrictions apply.

data within a system simultaneously. This leads to increase in data processed per unit time which is better known as throughput. A 16 stage pipeline as shown in Figure 1 is employed in this paper which means that each feistel round will act as one stage of pipelining. Registers are added after every round which act as output of the previous round for the current clock cycle and input for the next round in the next clock cycle. One register is also added after the final permutation which gives output only when the system is enabled. The system takes 16 clock cycle latency for the initial result and then it encrypts 64 bits of data per clock cycle moving forward. Each feistel round is taken care of with a

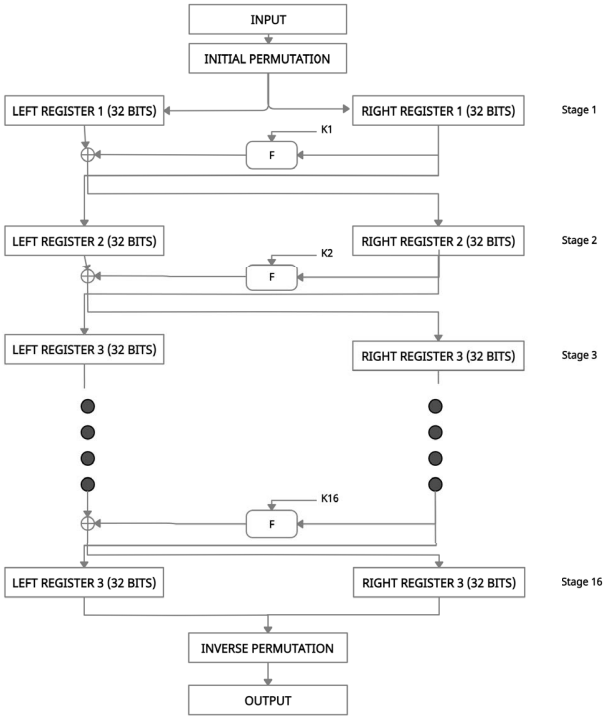


Fig. 1. Pipelining in DES

subkey which is obtained by left shift operations in the 56 bit new key obtained from the original key (64 bit) only by dropping off few bits. The procedure for generation of 16 sub keys is described in detail in [5]. The sub key generation process is also pipelined in accordance with the pipelining in the feistel rounds which can be seen in Figure 2. Each sub key is subject to be stored in registers after the corresponding left shift operation pertaining to the concerned round. These stored values are passed on to next stages to ensure a smooth pipelined design.

Both, the encrypting and decrypting blocks for DES are same as before but the only difference between the DES encrypt and DES decrypt schemes are the sequence in which the 16 sub-keys generated are fed to the Feistel rounds. The sequence in which sub keys are inserted in decryption is exactly reverse of the order in encryption. In non pipelined decryption of DES one may generate all the sub keys first and then apply them in reverse order i.e K16 to round 1, K15 to round 2 and so on. But K16 is generated only after all other sub keys are generated thereby disturbing the timing of the feistel rounds. So the traditional method of left shift operations to generate sub keys will not fit into the pipelined architecture of DES decryption. In order to generate sub keys in reverse order i.e K16, K15, K14 and so on, this paper performs right shift operations on the 56 bit initial key obtained from the user instead of left shift operations. Depending on the round the number of bits by which the keys have to be shifted will change. The number of bits shifted right in the decryption cycle is same as the number of bits shifted to the left in the encryption for the corresponding rounds with just round 1 being an exception where there is no shifting of the bits.

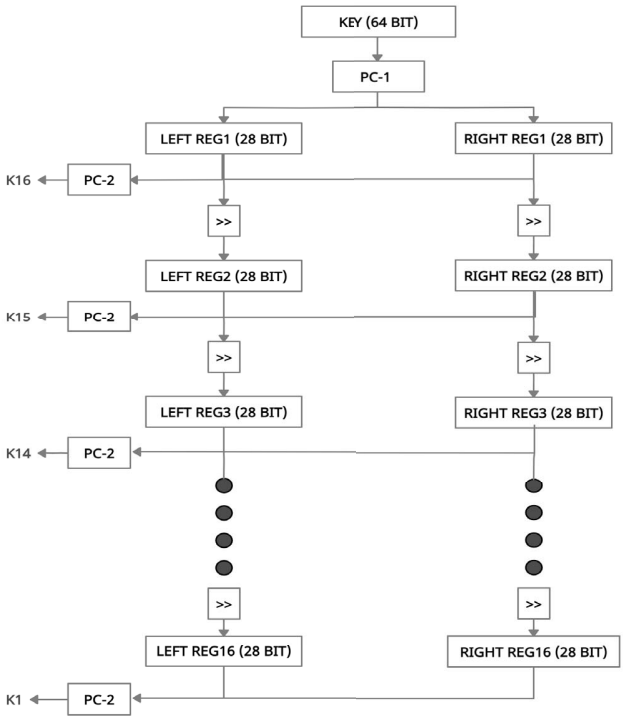


Fig. 2. Pipelined sub key generation in DES decryption

III. Pseudo-randomized Key generation

A. Linear Feedback Shift Registers

Linear Feedback Shift Registers are widely used in cryptographic applications, digital broadcasting , generation of counters, VLSI testing etc. A linear feedback shift register (LFSR) is a configuration of shift registers in which the input bit has a linear dependence on the previous stage [8] [9]. The XOR operation is the most widely used linear function. The input bit of the Linear Feedback Shift Register is the resultant of the XOR operation of few selected bits of the previous state values of the overall LFSR. Firstly the LFSR is loaded with an initial seed value which can be anything except all 0s. Then the LFSR is clocked and enabled i.e. when the input

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:58:51 UTC from IEEE Xplore. Restrictions apply.

message is changed and the edge (positive or negative) of the clock is encountered, it starts shifting data sequentially and performs XOR operation with the current value at the flip flops specified resulting in a pseudo random pattern of 1s and 0s. The few selected bits which are used in XOR operation define the characteristic polynomial of the LFSR. For example in Figure 3 the characteristic polynomial is

$$\phi(x) = x^4 + x + 1 \tag{1}$$

x^4 + 1 term indicates the presence of the loop and rest of the terms define the bits which are involved in the XOR operation (D1 is involved here so x^1 is the additional term). For LFSRs with large number of bits, the characteristic polynomial proves to be useful for representation and analysis as well. The bigger the LFSR is used, bigger is the length of pseudo random sequence thereby making it highly unpredictable and very secure. The pseudo randomizer used in this paper is achieved with the help of a 56 bit Linear Feedback Shift Register. The characteristic polynomial for the LFSR used in this paper is

$$\phi(x) = x^{56} + x^{48} + x^{36} + x^{24} + x^{12} + 1 \tag{2}$$

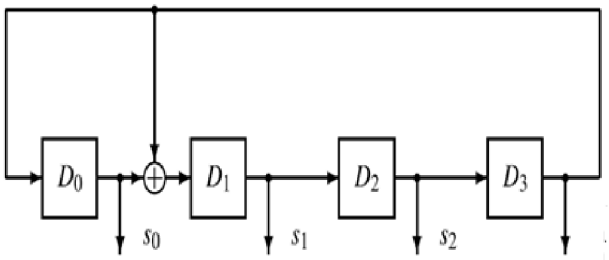


Fig. 3. A 4 bit LFSR

B. Implementation of the modified key generator

The aim of randomizing the subkeys is to increase the security. If the key is kept same for a stream of message bits then the security can be breached easily. In order to make these attacks by hackers face difficulty, it is ensured that for each different 64 bit message the key must be different and random. The original 56 bit key and the output of 56 bit LFSR are put under a specific function in order to get a new and random key from which subkeys are generated for Feistel rounds. The process is described below and is also depicted in Figure 4.

- Obtain the 56 bit key (Koriginal) by dropping off 8 select bits from 64 bit key(Kuser)
- Knew,i= f(Koriginal, Ri) (where Ri is ith 56 bit random number generated by LFSR)
- The function f used applies the following conditions for output:
  - If Ri[n] is 1 then complement the nth bit of Koriginal
  - If Ri[n] is 0 then retain the nth bit of Koriginal
- The obtained Knew,i is used to generate the subkeys for the feistel rounds

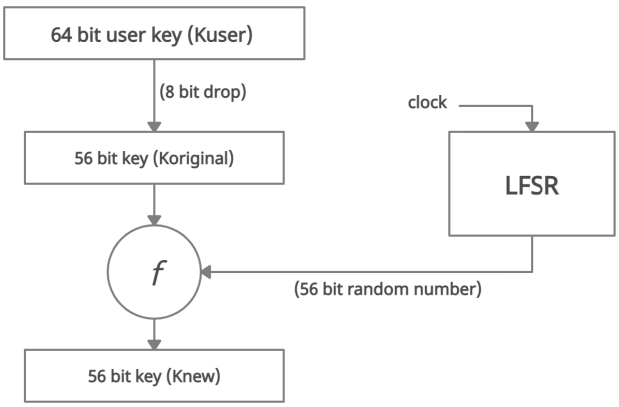


Fig. 4. Key randomization process

The value of LFSR changes with every change detected in message. Usually a stream of bits is sent over a communication system and hence after every clock the LFSR value is subject to change. The change in LFSR after every clock cycle also changes the new key to be used which in turn reduces the correlation between two consecutive encryption cycles thus making it more secure. The LFSR operation can be understood with the help of flowchart shown in Figure 5 and pseudocode discussed below:

Always @(posedge clock)
If(reg1 == plaintext) // plaintext is same as previous one
Lfsr_enable ⇐ 0; // lfsr value remains same
Else begin // plaintext changes
Lfsr_enable ⇐ 1; // lfsr value changes
Reg1 ⇐ plaintext; // new value of plaintext loaded for future comparisons

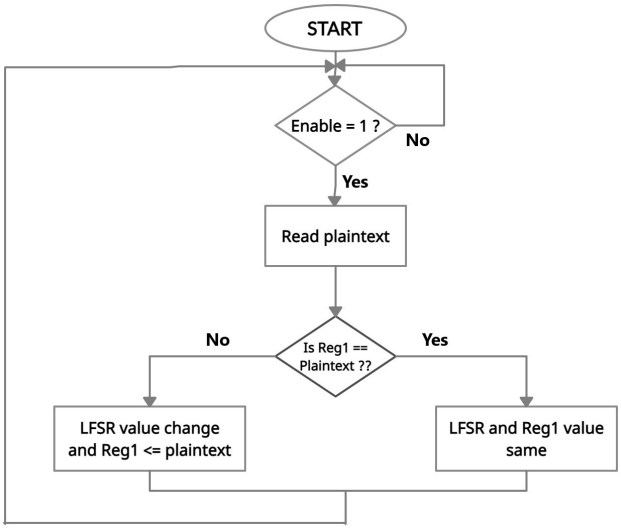


Fig. 5. Flowchart for LFSR operation.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:58:51 UTC from IEEE Xplore. Restrictions apply.

IV. Pseudo-Randomized Pipelined TDES

A. Encryption and Decryption

The TDES encryption process is pretty simple once DES algorithm is thoroughly understood. It is the connection of a DES encrypt block followed by a DES decrypt block then another DES encrypt block in series. With the help of structural coding practices in Verilog one can easily instantiate the DES blocks three times and connect them using wires. The objetive behind using three DES blocks is to enhance the security by increasing the key length and overcome the drawback of short key length in single DES system. The input keys are also connected to the three blocks using the key distribution process shown in Figure 6. The DES blocks which are instantiated here are the ones with pseudo randomized key generation block. The middle block uses an inverted primary clock (negative edge triggered with respect to primary clock) whereas the start and the end block use primary clock only(positive edge triggered with respect to primary clock). This is done to provide half clock cycle span for modified key function to generate its values correctly. Therefore the TDES discussed in this paper gives its first output after 16+ 0.5 +16 + 0.5 +16 = 49 clock cycles and thereafter gives the encrypted output after every clock cycle. In TDES there are three DES blocks, so in order to further improve the security user can feed different seed values to the three LFSRs to reduce the correlation between the intermediate processes.The TDES decrypting process is just the opposite of the TDES encrypting process. Therefore, a DES decrypt block followed by a DES encrypt block then another DES decrypt block are used for decryption. Since TDES is also a symmetric key algorithm it uses the same keys as the encryption process.

B. Key Management

A 192-bit key is provided by the user which is split up into 3 keys whose sizes are 64 bit each which act as the keys for each of the 3 DES blocks. Once again the problem encountered here is the lack of orderliness. To get over it the method so employed is key buffering i.e. use of buffers to synchronize the sets of 64 bit keys in the orientation as intended so that correct result is obtained. The first 64 bit key is fed into the first DES block as soon as it is given by the user. The next 64 key bit key is given to the second DES block after a 16 clock cycles buffer so that the first block can complete its computation since it takes an initial 16 clock latency. Lastly the third 64 bit key is fed to the third DES block after 33 clock cycles buffer so that the computations of first and second block is already complete when the third block's computation initiates. This ensures that if key changes at a particular instant, then the data which is already being processed is not fed with wrong keys and hence the results are not affected. Figure 6 shows the implementation of the TDES algorithm proposed in this paper.

V. Results and Discussion

A. Simulation Results

The simulations have been carried out in Xilinx ISE simulator better known as ISIM and are shown in Figure 7. The

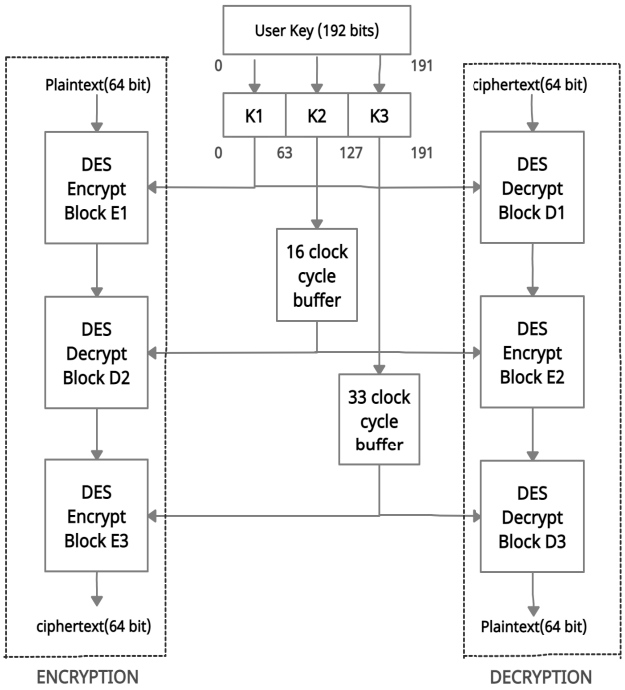


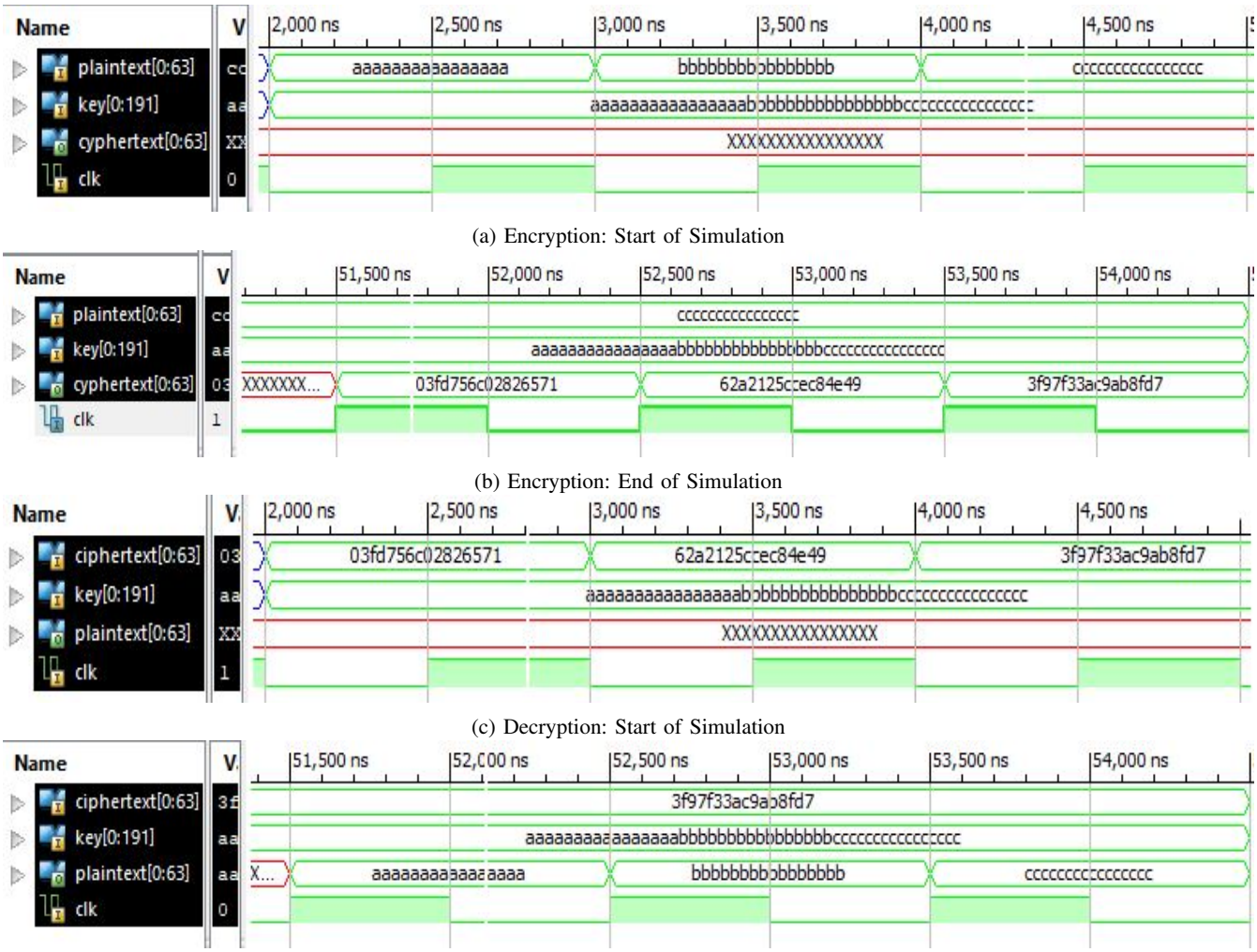
Fig. 6. TDES Block Diagram

message and master key are provided to the proposed design by the user in hexadecimal format at the start of the simulation. Three plaintexts are applied consecutively after each clock cycle. The sequence of plaintext applied is aaaaaaaaaaaaaaaa (1st clock) ,bbbbbbbbbbbbbbbb(2nd clock) , ccccccccccccccc(3rd clock). The 192 bit key is also fed which is "aaaaaaaaaaaaaaaa-bbbbbbbbbbbbbbbb-ccccccccccccccc". These plaintexts will be encrypted as per the TDES algorithm proposed in this paper. The first encrypted ciphertext appears after an initial latency of 49 clock pulses there after encryption occurs after each clock pulse. The sequence of encrypted ciphertext is 03fd756c02826571 (49th clock) , 62a2125ccec84e49(50th clock), 3f97f33ac9ab8fd7(51st clock). In order to verify decryption scheme of the proposed TDES algorithm, the ciphertext sequence is fed into the decrypt module in the same order i.e 03fd756c02826571 (1st clock) , 62a2125ccec84e49(2nd clock), 3f97f33ac9ab8fd7(3rd clock)with the same key which is fed in encryption. After 49 clock pulses the initial sequence of plaintext applied at the encryption is obtained back at the output i.e aaaaaaaaaaaaaaaa (1st clock) ,bbbbbbbbbbbbbbbb(2nd clock) , ccccccccccccccc(3rd clock).

B. Synthesis Results

In order to observe the improvements attained in the design, the non-pipelined and pipelined implementations of TDES were also done on the same device i.e xc7vx330t-3ffg1157. It is evident from Table I that both the pipelined and pseudo randomized TDES implementations are faster than the non

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:58:51 UTC from IEEE Xplore. Restrictions apply.



(a) Encryption: Start of Simulation

(b) Encryption: End of Simulation

(c) Decryption: Start of Simulation

(d) Decryption: End of Simulation

Fig. 7. Simulation Results

pipelined implementation but the resources used are more in the pipelined versions. Between pseudo randomized TDES and simple pipelined TDES, the latter is faster by a marginal amount which can be afforded as pseudo randomized TDES offers much better security. The resources used in all the three implementations are also well within the limits of the device used and further changes, if required can be made in the design without any difficulty.

TABLE I. Comparison between Non pipelined , Pipelined and pseudo randomized TDES

Type of Encrytion Used	Non-pipelined TDES	Pipelined TDES	Pseudo-randomized pipelined TDES
Maximum Frequency(MHz)	11.89	640.16	608.22
Maximum Throughput(Mbps)	761	40,970	38,926
Slice Registers Used	0	6120	6853
Slice LUTs used	8322	5377	6220
Delay(ns)	84.1	1.56	1.644

TABLE II. Comparison with previous works

Device Used	Authors	Maximum Frequency (MHz)	Maximum Throughput (Mbps)
xc4vlx25-12-ff676 (Virtex-4)	Fang Ren et al. [10]	215.165	860.66
	Proposed work	235.673	15083.072
xc7a100t-3csg324 (Artix-7)	Guruprasad S.P et al. [11]	348.947	797.593
	Proposed Work	391.654	25065.856
xc6vlx240t-3ff1156 (Virtex-6)	Saeid Taherkhani et al. [12]	294.031	18817.98
	Proposed work	566.370	36247.68

In order to reassure the fastness of the pseudo randomized TDES algorithm discussed in this paper, the speed parameters of the proposed work were compared with the previous works discussed in [10] [11] and [12]. The proposed design was implemented on the same device with which these works have been carried out and the results are shown in Table II. It is observed that proposed design is faster than the previous works which is depicted in Figure 8. It is also seen that the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:58:51 UTC from IEEE Xplore. Restrictions apply.

work presented in this paper gives tremendous increase in throughput (processes more data in less amount of time) as compared to all the works which have been taken for reference as can be seen in Figure 9. This is because the proposed work will process each batch of 64 bit data after every single clock cycle whereas the previous works [10] and [11] take multiple clock cycles to process a single batch of 64 bit data. It is also worthy to note that all these improvements in the proposed design are coming along with improvement in security due to the pseudo randomized key generation.

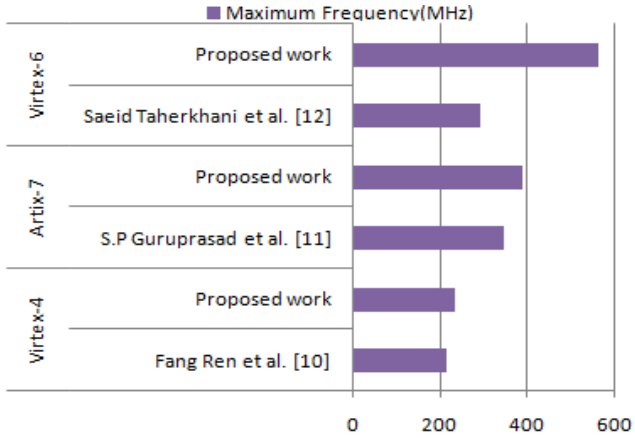


Fig. 8. Graphical representation of comparison of maximum frequency with previous works

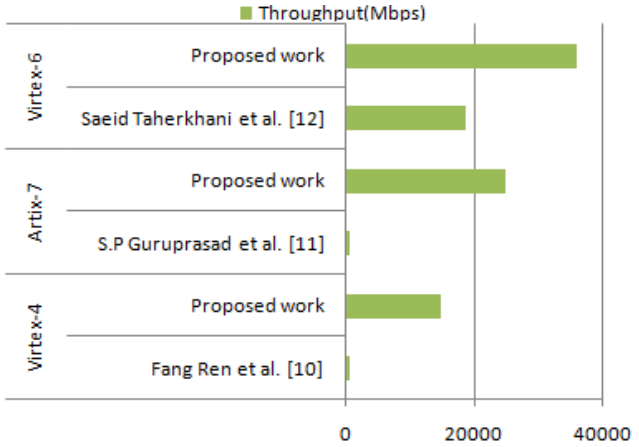


Fig. 9. Graphical representation of comparison of maximum throughput with previous works

VI. Conclusion

A highly efficient and secure implementation of TDES algorithm is presented in this paper. The proposed work offers enhanced speeds with the help of the pipelining explained in the previous sections and is way secure than any traditional TDES system because of its modified key generation which employs a pseudo randomizer (with the help of an LFSR) which makes the encrypted output random thereby increasing the sense of security. The proposed system is faster and provides enhanced throughput when compared to previous works [10] [11] and [12] on respective devices. The design is implemented on the device xc7vx330t-3ffg1157 (Virtex – 7) using which maximum frequency of operation achieved is 608.224 MHz and at the maximum throughput so achieved is 38.926 Gbps. This variant of TDES encryption can be used in various real time applications like electronic payment systems, SMS encryption, cloud security, smart card applications etc. because of its increased sense of security and speed.

References

[1] M. E. Smid, D. K. Branstad, “Data Encryption Standard: past and future,” in Proceedings of the IEEE, vol. 76, no. 5, pp. 550-559, May 1988
[2] R. Bhanot, R. Hans, “A Review and Comparative Analysis of Various Encryption Algorithms,” International journal of security and its applications, vol. 9, no. 4, 2015, pp. 289-306
[3] A. Nadeem, M. Y. Javed, “A Performance Comparison of Data Encryption Algorithms,” 2005 International Conference on Information and Communication Technologies, Karachi, Pakistan, 2005, pp. 84-89.
[4] E. Barker, N. Mouha, “Recommendation for the Triple Data Encryption Algorithm (TDEA) Block Cipher,” NIST Special Publication, National Institute of Standards and Technology, Gaithersburg, 2017, pp. 800-867.
[5] Data Encryption Standard (DES), National Institute of Standards and Technology, Gaithersburg, 1999
[6] E. D. Rosal, S. Kumar, “A Fast FPGA Implementation for Triple DES Encryption Scheme,” Circuits and Systems, vol. 8, 2017, pp. 237-246.
[7] S. Kilts, “Advanced FPGA Design: Architecture, Implementation and Optimization”, Wiley-IEEE Press, June 2007, pp. 1-16.
[8] A. K. Panda, P. Rajput, B. Shukla, “FPGA Implementation of 8, 16 and 32 Bit LFSR with Maximum Length Feedback Polynomial Using VHDL,” 2012 International Conference on Communication Systems and Network Technologies, Rajkot, Gujarat, India, 2012, pp. 769-773.
[9] A. Ajane, P. M. Furth, E. E. Johnson, R. L. Subramanyam, “Comparison of binary and LFSR counters and efficient LFSR decoding algorithm,” IEEE 54th International Midwest Symposium on Circuits and Systems (MWSCAS), Seoul, Korea (South), 2011, pp. 1-4.
[10] F. Ren, L. Chen., T. Zhang, “3DES Implementation Based on FPGA,” Emerging Research in Web Information Systems and Mining, 2011, pp. 218–224.
[11] S. P. Guruprasad, B. S. Chandrasekar, “An evaluation framework for security algorithms performance realization on FPGA,” 2018 IEEE International Conference on Current Trends in Advanced Computing (ICCTAC), Bangalore, India, 2018, pp. 1-6.
[12] S. Taherkhani, E. Ever, O. Gemikonakli, “Implementation of Non-Pipelined and Pipelined Data Encryption Standard (DES) Using Xilinx Virtex-6 FPGA Technology,” 2010 10th IEEE International Conference on Computer and Information Technology, Bradford, UK, 2010, pp. 1257-1262.
[13] K. Chitra, V. P. Venkatesan. ”An antiquity to the contemporary of secret sharing scheme,” Journal of Innovative Image Processing (JIIP) ,vol. 2, no. 1, 2020, pp. 1-13.
[14] I. R. S. Reddy and G. Murali, ”A novel triple DES to enhance E-governance security,” 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), Chennai, India, 2017, pp. 2443-2446
[15] S. Shrivastava, S. Palnitkar, S. Pathak, M. Shinde, ”Secure Encryption of data using Symmetric and Asymmetric Cryptographic Algorithm,” Mukt Shabd Journal, vol. 9, no. 6, June 2020, pp. 1100-1109
[16] S. Kaur and L. Jain, “A Hybrid Cryptographic Scheme for Improving Cloud Security Using ECC and TDES Algorithms”, Current Journal of Applied Science and Technology, vol. 39, no. 47, pp. 27-34, December 2020.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:58:51 UTC from IEEE Xplore. Restrictions apply.

2021 2nd International Conference for Emerging Technology (INCET)
Belgaum, India. May 21-23, 2021

From Classical to Quantum: A Review of Recent Progress in Reinforcement Learning

Kunal Kashyap
Department of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
kunalkashyap1107@gmail.com

Lalit
Department of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
lalit31999@gmail.com

Daksh Shah
Department of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
dakshshah0605@gmail.com

Lokesh Gautam
Department of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
lokeshgautam21@gmail.com

2021 2nd International Conference for Emerging Technology (INCET) | 978-1-7281-7029-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/INCET51464.2021.9456218

***Abstract*—Standard machine learning algorithms have taken shape well due to increase in computational power. There is a special focus on reinforcement learning, which offers solutions to concepts like autonomous vehicles and robotics, among others. On the other hand, research in the fields of quantum computing and associated machine learning has catapulted in the last few years. This work analytically compares and contrasts between the classical and quantum approaches to two reinforcement learning problems, and tries to put focus on future works to advance both fields further.**

Keywords—reinforcement learning, quantum computing, q-learning, variational quantum circuits

I. INTRODUCTION

Reinforcement learning (RL) refers to a class of machine learning problems which make use of learning the steps for maximizing a numerical reward, specifically by learning about situations and actions and the possible mappings between each [1]. These are most distinguished by being crucially closed-loop, not being informed about possible actions at any state, and how a particular action can lead out to multiple subpaths over repeated experiments. Some intuitive examples may include a game of chess, an adaptive controller at a refinery, and a mobile robot.

One can segregate an RL system into four main sub-components (excluding the agent and the environment):

- *Policy:* It describes, at a given time, the learning agent's way of behaving. Psychologically, it can be termed as "a set of stimulus response rules or associations". Often called the core of an RL agent, it alone is sufficient to ascertain the agent's behavior with respect to the environment.
- *Reward signal:* It is essentially the goal in an RL problem. The agent is being incentivized to maximize the cumulative reward received during each iteration. In a way, it describes the positive and negative events for the agent.
- *Value function:* For an arbitrary state, its *value* is the total amount of reward the agent can hope to collect over the future, starting from that state. Thus, the value function signifies the long-term desirability of states, where rewards determine the immediate and intrinsic desirability.
- *Model:* Used in some RL systems, it can mimic the behavior of the environment. In other words, it can allow reasonings to be furnished on how the environment will behave.

On the other hand, the field of quantum machine learning (QML), a subset of quantum computing [2] has been a major focus of research in recent years, partly due to development of quantum computers by prolific technology firms such as IBM [3] and Google [4]. The mix of quantum computing, which offers infinite superpositions of quantum bits (qubits) and quantum information transfer (instantaneous transfer), with learning algorithms offers possibilities of predictions more accurate than ever [5].

This work is an effort to analytically compare and contrast between two approaches to the problem of machine learning: the first being Q-learning, a standard model-free technique for reinforcement learning, and the second being an RL system made using variational quantum circuits on noisy intermediate-scale quantum (NISQ) devices [6], [7]. It also endeavors to comment on future trends based on the approaches presented, and hopefully guide research in the right direction.

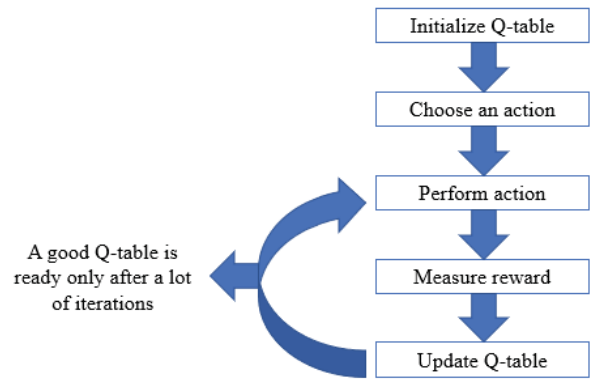


Fig. 1. Flowchart showing the working of Q-learning algorithm.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:49:54 UTC from IEEE Xplore. Restrictions apply.

The final Q-table generated is good enough for testing use only after a considerable number of iterations.

II. Q-Learning

Q-Learning is one of the recent reinforcement learning methods which is used for learning the optimal policy in a Markov Decision Process [1]. In reinforcement learning, after every action, a state is reached and accordingly the next step has to be taken. This state-action pair is taken care by the Q-values. The main desideratum of the Q-learning is to maximise the total reward or the return after every action taken by the algorithm.

The Q-function for a particular policy takes in a state and action pair, and gives out an expected reward from performing that particular action in that particular state, following the particular policy all along. It is based on the following Bellman's Equation:

$$Q^{\pi}(s_t,a_t)=E[R_{t+1}+\gamma R_{t+2}+\gamma^2 R_{t+3}+\ldots][s_t,a_t], \quad (1)$$

where the given state and action pair is multiplied with the expected discounted cumulative reward so that Q-values are created for that pair.

After every action, there is an iteration and the Q-values are updated. This iteration process continues till we get an optimal set of Q-values in the Q-Table. This table gives the idea of the Q-values with respect to various actions taken and returns received. Initially all the values in Q-table are 0 and then after every iteration, the values are updated.

The flow diagram for the functioning of the Q-Algorithm is shown in Fig.1. When we are able to get the maximum return, a good Q-table is achieved.

The algorithm works on the concept of Exploration and Exploitation. When it is just the start of the process and there is no idea of current surroundings, the process of exploration is executed so as to study the environment. Once the reinforcement learning takes place and the idea of return on future action is known, exploitation has to be done. A good Q-learning algorithm makes a balanced use of exploration as well as exploitation.

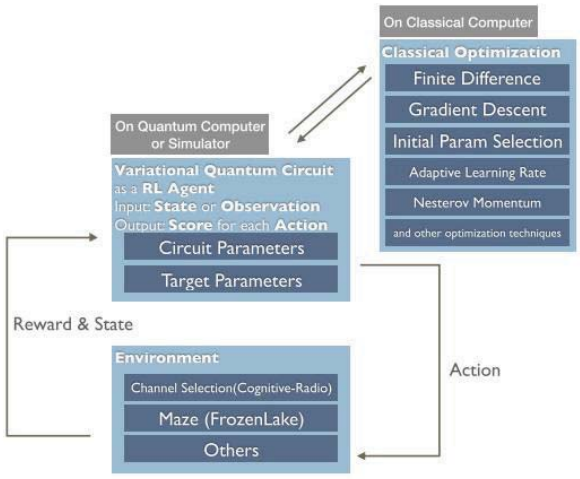


Fig. 2. High-level view of variational quantum circuits for deep reinforcement learning [8].

A hybrid quantum-classical approach is described to solve the RL problem. The framework suggested in [8] can be expected to solve more complex tasks when quantum machines are available in a larger scale.

III. Variational Quantum Circuits for RL

Reference [8] was the first to reshape concepts more akin to the classical approach to deep reinforcement learning, e.g. experience replay and target network, into an actualization of variational quantum circuits. Also, a computational basis encoding scheme was used to reduce the number of parameters controlling the model required for deep reinforcement learning in the classical scenario.

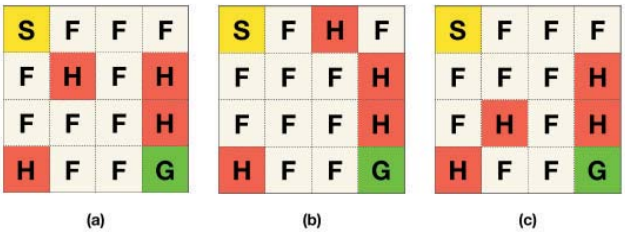


Fig. 3. Frozen-lake environment [8].

Frozen-lake environment is a simple maze environment in openAI Gym [9]. The RL agent is supposed to go from the start position (marked S) to the goal position (marked G) one step at a time while avoiding holes (marked H).

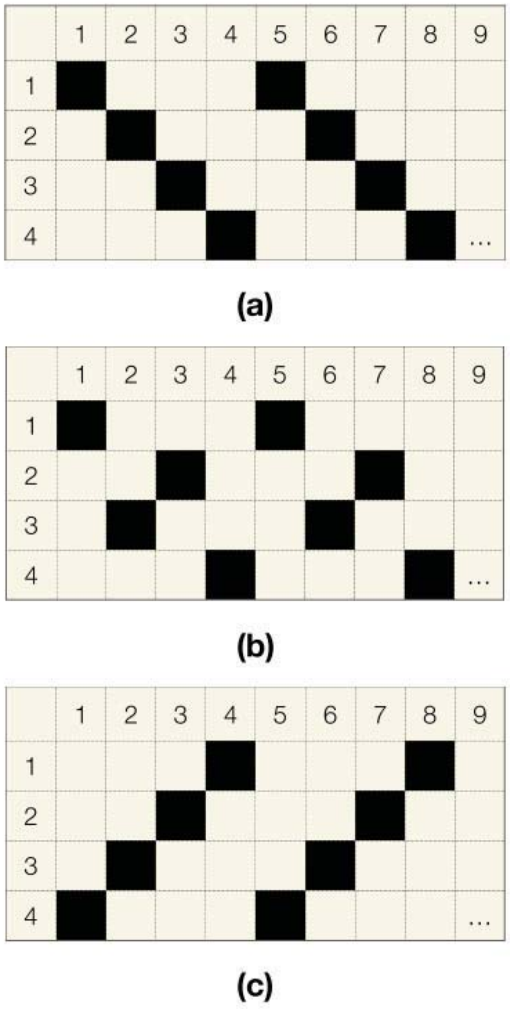


Fig. 4. Cognitive-radio environment [8].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:49:54 UTC from IEEE Xplore. Restrictions apply.

Cognitive-radio environment is used to simulate a real-world application. The agent is supposed to select a free channel not occupied or in use by a primary user (shaded black). To avoid complexity, three simple periodic patterns were utilized.

Two testing environments were considered for the ensuing experiments: frozen-lake and cognitive-radio. The frozen-lake environment is a fairly simple and commonly tested example in reinforcement learning (see Fig.3), and feasible simulations are possible with available quantum machines if the problem size dimension is not too large (with reasonable time consumption for the experiment). Using this environment, it was also established that quantum circuits have the capability to learn a policy, or a process where decisions are to be made in sequence. The cognitive-radio environment (see Fig.4), while equivalent to the frozen-lake environment in terms of complexity, helped establish the capability of these circuits to execute a real-world application.

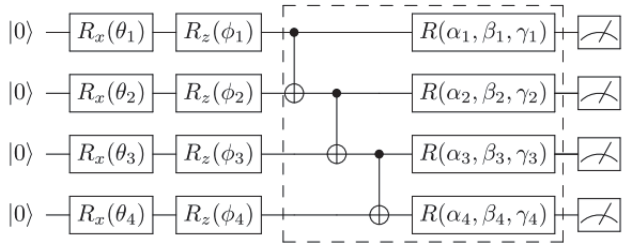


Fig. 5. VQ-DQN proposed for the frozen-lake testing environment [8].

Rotation about the *x*-axis and *z*-axis is provided by the single-qubit gates at the beginning of the circuit. The CNOT gates entangle the quantum states. The general single-qubit unitary gates provide the iterative optimization.

The variational quantum circuit is an approach adopting the strengths of both classical and quantum aspects of computing. It has tunable parameters optimized in an iterative manner by a classical computer, which are equivalent to weights in an artificial neural network. This approach is flexible with respect to the depth of circuits and somewhat resistant to noise. Therefore, by combining a variational quantum circuit, QML algorithms can bypass complex quantum errors present in current devices, even when quantum error correction and fault-tolerant computation is not available in current quantum machines.

The reinforcement learning agent (Variational Quantum Deep Q-Network – VQ-DQN) proposed was divided into quantum and classical parts (see Fig.2), the former outputting measurement results based on two sets of parameters, determining possible actions to take. The latter part focuses on optimizing the circuit by deciding the new values of parameters. The same circuit architecture was used to construct two sets of parameters. Optimization was achieved by calculating gradients of expectation values of quantum measurements using the unchanged circuit architecture and slenderly different parameters. Each state was further encoded with computational basis encoding, i.e., each state was represented in binary and then corresponding quantum state was calculated from that result. The encoding was done using the single-qubit unitary rotation method.

The variational quantum circuit is flexible in circuit depth, which means that a shallow circuit can still approximately achieve goals set by the solution space, but a deep, more sophisticated circuit may perform better. Thus, a strongly entangling low-depth circuit was used twice for both testing environments. The number of qubits used in the aforementioned circuit can be modified to fit the problem posed by the respective testing environment as well as the quantum computer or simulator.

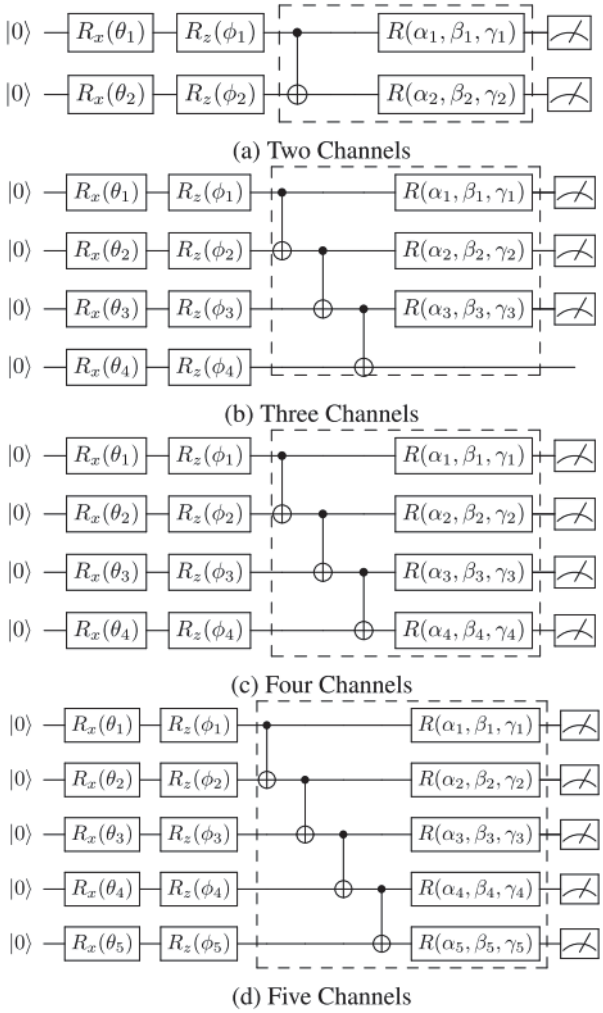


Fig. 6. VQ-DQN proposed for the cognitive-radio environment [8].

A four-qubit circuit was proposed for the three-channel case because the number of environment states (i.e., 9) is greater than the number of computational basis states (i.e., 8). The networks can be seen as an extension for that proposed for the frozen-lake environment.

The VQ-DQN for both testing environments can be seen in Fig.5 and Fig.6. For an *n*-qubit quantum circuit, the measurement output was taken to be the ensemble average (i.e., expectation) value for a single qubit. The output for *n* qubits was stored in an *n*-tuple. The index of the above measurement output port/wire which has the largest expectation value was used as the action for the circuit. The subsequent action that the agent then chooses was determined by the corresponding output for a qubit, and not some random outcome in each single run of the experiment. Therefore, the discrete probability distribution can be learned in O(*n*) time complexity. Quantum circuits also provide the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:49:54 UTC from IEEE Xplore. Restrictions apply.

TABLE I. COMPARISON BETWEEN CLASSICAL Q-LEARNING AND PROPOSED VQ-DQN FOR FROZEN-LAKE AND COGNITIVE-RADIO TESTING ENVIRONMENTS.

Environment	2-Channels	3-Channels	4-Channels	5-Channels	Frozen-Lake
Q-Learning	$2 \times 2 \times 2$	$3 \times 3 \times 3$	$4 \times 4 \times 4$	$5 \times 5 \times 5$	$4 \times 4 \times 4$
VQ-DQN	$2 \times (3 \times 2 + 1)$	$3 \times (3 \times 2 + 1)$	$4 \times (3 \times 2 + 1)$	$5 \times (3 \times 2 + 1)$	$4 \times (3 \times 2 + 1)$

flexibility of parallel measurement across each qubit, which means that the number of repeated experiments can be a fixed number and not dependent on n. Moreover, because the focus is on the qubit having the highest expectation value, the algorithm proposed was comparatively potent against noise and errors present in real quantum machines.

Two kinds of simulations were performed as part of the research: one being with software packages PennyLane [10] and PyTorch [11] (for simulation acceleration), the other with IBM Qiskit-Aer simulation software as the backend to incorporate noise from a real quantum computer.

The lack of quantum error correction in current NISQ machines implies that they are highly unsuitable for deep quantum circuit architectures. A rough corollary can be that when the number of quantum gates increases in a circuit, there will exist more errors. The circuit proposed in [8] thus cannot have parameters commensurate to classical neural networks, but was able to show comparable performance for the environments considered with fewer iterative parameters.

IV. COMPARISON BETWEEN APPROACHES

Table 1 shows for both testing environments, the number of parameters in classical Q-learning is of the form $n \times n \times n$, where n is a dimension of the environment [8]. This means that as the environment gets more complex, the number of parameters increases, making the RL problem more computationally expensive.

On the other hand, the VQ-DQN approach makes use of two-qubit CNOT gates and parameterized single-qubit unitary gates, leading to a total of $n \times (3 \times 2 + 1)$ parameters in its learning process. So, in general, the circuit can represent the n^2 states the testing environment requires with an n-qubit system, providing 2^n computational basis states in the process.

An exception is $n = 3$, where the number of states required is greater than the number of computation basis states. Here, the proposal was to build a four-qubit circuit but leave out the parameterization and measurement of the fourth qubit. A graphical comparison can be seen in Fig.7.

From Table 1 as well as Fig.7, the VQ-DQN approach (for the cognitive-radio environment) uses less parameters for cases with large number of channels than the classical Q-learning approach.

V. ANALYSIS

We saw from the two perspectives of solving the testing environments the efficiency of variational quantum circuits, which implies that this approach can be applied on near-term NISQ machines while bypassing difficulties intrinsic to quantum computing such as quantum error correction.

Current research in the quantum computing field is limited, mostly due to the level of complexity involved in building and maintaining a quantum computation device. In spite of this limitation, considerable progress has been made, surpassing existing applications in terms of implementations and creating new ones in the process [12]–[15].

Future scope with the quantum approach to reinforcement learning can be seen with everyday devices edge-connected with quantum machines to provide precise judgments and decisions for future steps. A notable mention in this scope is autonomous vehicles, which if connected to a quantum computer able to provide real-time actions may not ever have to deal with choices such as the runaway trolley paradox. This is, of course, only feasible if and when quantum machines are built to scale.

In conclusion, the problems faced with quantum computing at present is because conclusive research on this topic is only at its nascent stage, so we can expect big disruptions in the field in the near future.

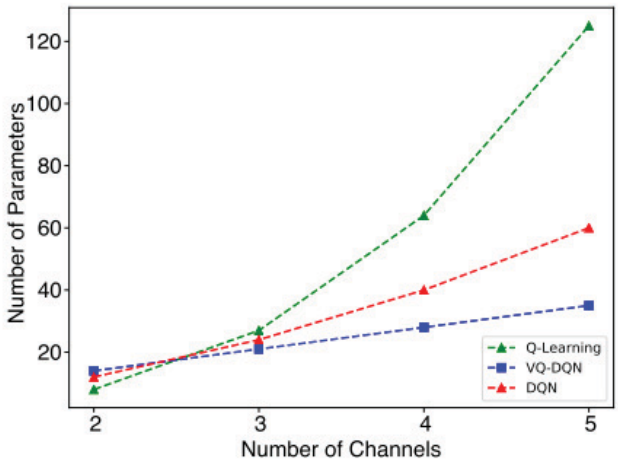


Fig. 7. Comparison of number of parameters required for solving the n-channel cognitive-radio testing environment [8].

The VQ-DQN approach is able to compute more efficiently with lesser number of parameters as the number of channels increases.

ACKNOWLEDGMENT

K. K., L. and D. S. thank Srijan Saxena and Sudhanshu Kumar for constructive discussion and great support.

REFERENCES

[1] P. R. Montague, "Reinforcement Learning: An Introduction, by Sutton, R.S. and Barto, A.G.," *Trends Cogn. Sci.*, vol. 3, no. 9, 1999, doi: 10.1016/s1364-6613(99)01331-5.

[2] IBM, "What is quantum computing ?," *IBM*, 2020.

[3] IBM, "IBM Q Experience," *Quantumexperience.Ng.Bluemix.Net*, 2019. .

[4] F. Arute *et al.*, "Quantum supremacy using a programmable superconducting processor," *Nature*, vol. 574, no. 7779, 2019, doi: 10.1038/s41586-019-1666-5.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:49:54 UTC from IEEE Xplore. Restrictions apply.

[5] D. Dong, C. Chen, H. Li, and T. J. Tarn, "Quantum reinforcement learning," *IEEE Trans. Syst. Man, Cybern. Part B Cybern.*, vol. 38, no. 5, pp. 1207–1220, 2008, doi: 10.1109/TSMCB.2008.925743.

[6] Y. Du, M.-H. Hsieh, T. Liu, and D. Tao, "Expressive power of parametrized quantum circuits," *Phys. Rev. Res.*, vol. 2, no. 3, 2020, doi: 10.1103/physrevresearch.2.033125.

[7] J. Preskill, "Quantum computing in the NISQ era and beyond," *Quantum*, vol. 2, 2018, doi: 10.22331/q-2018-08-06-79.

[8] S. Y. C. Chen, C. H. H. Yang, J. Qi, P. Y. Chen, X. Ma, and H. S. Goan, "Variational Quantum Circuits for Deep Reinforcement Learning," *IEEE Access*, vol. 8, pp. 141007–141024, 2020, doi: 10.1109/ACCESS.2020.3010470.

[9] G. Brockman *et al.*, "OpenAI Gym," pp. 1–4, 2016, [Online]. Available: http://arxiv.org/abs/1606.01540.

[10] V. Bergholm *et al.*, "PennyLane: Automatic differentiation of hybrid quantumclassical computations," *arXiv*. 2018.

[11] S. Zagoruyko *et al.*, "A multipath network for object detection," in *British Machine Vision Conference 2016, BMVC 2016*, 2016, vol. 2016-September, doi: 10.5244/C.30.15.

[12] A. D. Corcoles *et al.*, "Challenges and Opportunities of Near-Term Quantum Computing Systems," *Proc. IEEE*, vol. 108, no. 8, 2020, doi: 10.1109/JPROC.2019.2954005.

[13] A. Ajagekar and F. You, "Quantum computing for energy systems optimization: Challenges and opportunities," *Energy*, vol. 179, 2019, doi: 10.1016/j.energy.2019.04.186.

[14] C. Outeiral, M. Strahm, J. Shi, G. M. Morris, S. C. Benjamin, and C. M. Deane, "The prospects of quantum computing in computational molecular biology," *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 11, no. 1. 2021, doi: 10.1002/wcms.1481.

[15] Y. Cao *et al.*, "Quantum Chemistry in the Age of Quantum Computing," *Chemical Reviews*, vol. 119, no. 19. 2019, doi: 10.1021/acs.chemrev.8b00803.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:49:54 UTC from IEEE Xplore. Restrictions apply.

Progress In Electromagnetics Research Letters, Vol. 98, 41–48, 2021

Gain Enhancement Using Modified Circular Loop FSS Loaded with Slot Antenna for Sub-6 GHz 5G Application

Anubhav Kumar[1, *], Asok De[2], and Rakesh K. Jain[1]

Abstract—In this paper, a modified circular loop Frequency Selective Surface (FSS) loaded on a slot antenna is proposed for sub-6 GHz 5G applications. The proposed FSS reduces the resonant frequency to lower bands of conventional circular FSS without change in its size. The operating bandwidth (−10 dB) of the proposed antenna with polarization-insensitive single-layer FSS varies from 3.6 GHz to 6.1 GHz with an average gain of 7–7.5 dB and a maximum realized gain of 7.87 dB. An FSS superstrate is loaded onto a slot antenna to increase the realized gain up to 4 dB, where the FSS shows desirable electromagnetic wave reflection characteristics over operating bandwidth and can be used in 5G sub-6 GHz band applications.

1. INTRODUCTION

Frequency selective surface (FSS) structure is used for gain, bandwidth enhancement and grating lobe reduction in microstrip antenna, which can reflect or pass the EM waves coming from the antenna. Several gain enhancement techniques have been investigated to enhance the properties of microstrip antenna such as artificial magnetic conductor (AMC) [5], metamaterial [1, 6, 13], single [8, 10], and double-layer FSS [2, 7, 11]. In [1], a double split ring reflector based on mu-negative characteristic of the metasurface is used for gain enhancement. The dual FSS layers based on stopband reflector are loaded on a ultra-wideband (UWB) antenna, where a split ring in the top layer and the combination of Jerusalem cross (JC) FSS with a square loop in the bottom layer are used for gain enhancement [2]. A parabolic cylinder FSS reflector is accomplished on a dielectric resonator antenna (DRA) and monopole antenna to enhance the gain [3]. Due to the tremendous reflecting and absorbing property of FSS, multilayer FSS is also a popular approach for enhancing the parameters of a microstrip antenna. In [4], a multi-layer bandpass FSS placed above the antenna is used to enhance the gain. In [5], a triple band AMC structure is used below the antenna for gain enhancement where AMC represents zero reflection phases in the resonant frequencies. In [6], split-ring resonator (SRR) metamaterial is loaded on a DRA to enhance the gain with stable efficiency in the operating bandwidth. In [7], a double-layer passband FSS is used to influence the gain at ISM band placed over microstrip antenna. In [8], the single-layer passband substrate integrated waveguide (SIW)-FSS is loaded to enhance the gain of a microstrip antenna whereas in [9] an FSS combined with a DRA is used for gain enhancement. In [10], double-side single layer FSS incorporated with chip resistor is used for gain enhancement of a DRA in X-band. Similarly in [11, 12], a double-layer FSS reflector is designed for UWB applications, improving the gain effectively. In [13], a metamaterial SRR incorporated with the antenna in the same substrate is used to enhance the gain in two operating bands. In the proposed paper, a modified circular loop FSS is designed. The insensitive polarization FSS is analyzed for the gain improvement of a slot antenna based on microstrip-feed in sub-6 GHz operating bandwidth.

Received 11 March 2021, Accepted 24 May 2021, Scheduled 30 May 2021

* Corresponding author: Anubhav Kumar (rajput.anubhav@gmail.com).

[1] Department of Electronics and Communication Engineering, Shobhit Institute of Engineering and Technology, (Deemed to be University) Meerut, Uttar Pradesh, India. [2] Department of Electronics and Communication Engineering, Delhi Technological University (DTU), New Delhi, India.

2. ANTENNA DESIGN WITH EQUIVALENT CIRCUIT (EC) MODEL

The proposed slot antenna with fabricated hardware is shown in Fig. 1(a) where dimensions are as follows (in mm): $W1 = 2$, $W3 = 1.5$, $W4 = 0.75$, $W5 = 9$, $W6 = 12$, $W7 = 13.5$, $W8 = 1$, $Ls = 30$, $L1 = 2$, $L2 = 7$, $L3 = 14.5$, $Wf = 3$. FR-4 material is utilized to design the antenna with a 4.4 dielectric constant. The proposed antenna size is $30 \times 30 \times 1.6\,\text{mm}^3$, which is based on a microstrip-fed and slot based defected ground. In the first step as model-I, 15 mm length (Lf) of microstrip-fed line with 3 mm width (Wf) is used for achieving 50-ohm impedance where the length of the microstrip-fed is directly related to the operating bandwidth. A rectangular slot is inserted in the ground as Model-I in Fig. 1(b) that influences the impedance matching which results in impedance bandwidth (-10 dB) from 3.4 GHz to 4 GHz as depicted in Fig. 1(c). In the second step as model-II, the internal impedance of an antenna is influenced by the variation of the ground, when the size of the rectangular slot of the ground increases and tilted 45°, which enhances the impedance bandwidth from 3.5 GHz to 5.9 GHz. In the last step, inductance width increases with some corner defect in the ground therefore resonant frequency of antenna shifted from 4.8 GHz to 4.7 GHz where impedance bandwidth varies from 3.5 GHz to 6.1 GHz as depicted in Fig. 1(c) in Model-III. The EC model of the proposed antenna is developed and tuned with QUCS software as represented in Fig. 1(d). The $|S_{11}|$ bandwidth of EC model is matched with antenna bandwidth illustrated in Fig. 1(e), where lumped elements of feed line (Lf, Cf), slot in-ground (Lg, Cg), and resonant frequency (Rp, Lp and Cp) are $Lf = 2$ nH, $Cf = 0.63$ pF, $Lg = 3.7$ nH, $Cg = 1.577$ pF, $Rp = 49\,\Omega$, $Lp = 1.96$ nH, and $Cp = 1.03$ pF.

3. DESIGN OF THE FREQUENCY SELECTIVE SURFACE

3.1. Design of the Unit Cell

The proposed modified circular loop FSS is optimized for a 5G sub-6 GHz application with a stopband characteristic, which varies from 3 GHz to 7.8 GHz. The design steps of a unit cell FSS are illustrated in Fig. 2(a) and discussed as follows:

Step-I — The FSS stopband and reflection characteristics depend on unit cell periodicity (P) and the incidence angle (θ_o) of the EM waves coming from the antenna. The periodicity (P) can be calculated by the wavelength (λ) of the EM waves, and its incidence angle is represented in Equation (1) [17].

$$\mathbf{P} < \frac{\boldsymbol{\lambda}}{\mathbf{1} + \mathbf{sin}\boldsymbol{\theta}_\mathbf{o}} \tag{1}$$

When incident angle varies from 0° to 60°, the FSS unit cell size should not exceed 26.79 mm for a 6 GHz frequency by implementing Equation (1). Therefore, in the first step, the customized size of the circular unit cell is taken 14 mm. The ring FSS circumference will be equal to one wavelength, and the radius should be 10.16 mm for 4.7 GHz [14], but here the circular unit cell FSS radius is taken as 6 mm, keeping the compact size as depicted in Fig. 2(a). The 10 dB bandwidth varies from 5.2 GHz to 6.9 GHz and is represented in Fig. 2(b) with resonance frequency 6.1 GHz.

Step-II — In the second step, two strips of 0.5 mm width and 12 mm length (v) have been combined with a ring FSS (step-I) to increase the 10 dB impedance bandwidth and reduce the resonant frequency without increasing the size. The impedance bandwidth (-10 dB) of the hybrid unit cell is achieved from 3.6 GHz to 5.7 GHz with a resonant frequency of 4.7 GHz where stopband frequencies vary from 3 GHz to 7.8 GHz, as illustrated in step-II and its s-parameters (in dB) in Fig. 2(a) and Fig. 2(b). The reflection phase (S_{11}) of FSS unit cell decreases linearly with increasing frequency. The unit cell of the FSS is analyzed with the master-slave condition in x- and y-planes and Floquet-port in z-plane to investigate the stopband response in different angles and polarizations. FSS polarization stability is important and requires characteristics for practical applications. The resonant frequency of an FSS must be stable on the EM waves coming from different angles and polarizations so that it can properly control the grating lobes. The polarization instability of the hybrid structure unit cell is extracted in TE/TM polarization from 0° to 60°, and its $|S_{21}|$ is depicted in Fig. 2(c) and Fig. 2(d). Modified circular FSS is angular stable at a different angle and polarization at an operating frequency. The lumped element of a square loop FSS is discussed in [15]. The overall values of L and C of the EC model are extracted by Equation (2) and Equation (3). The factor $\pi/2$ in Bc and $\pi/4$ in XL is used because of the circular loop structure

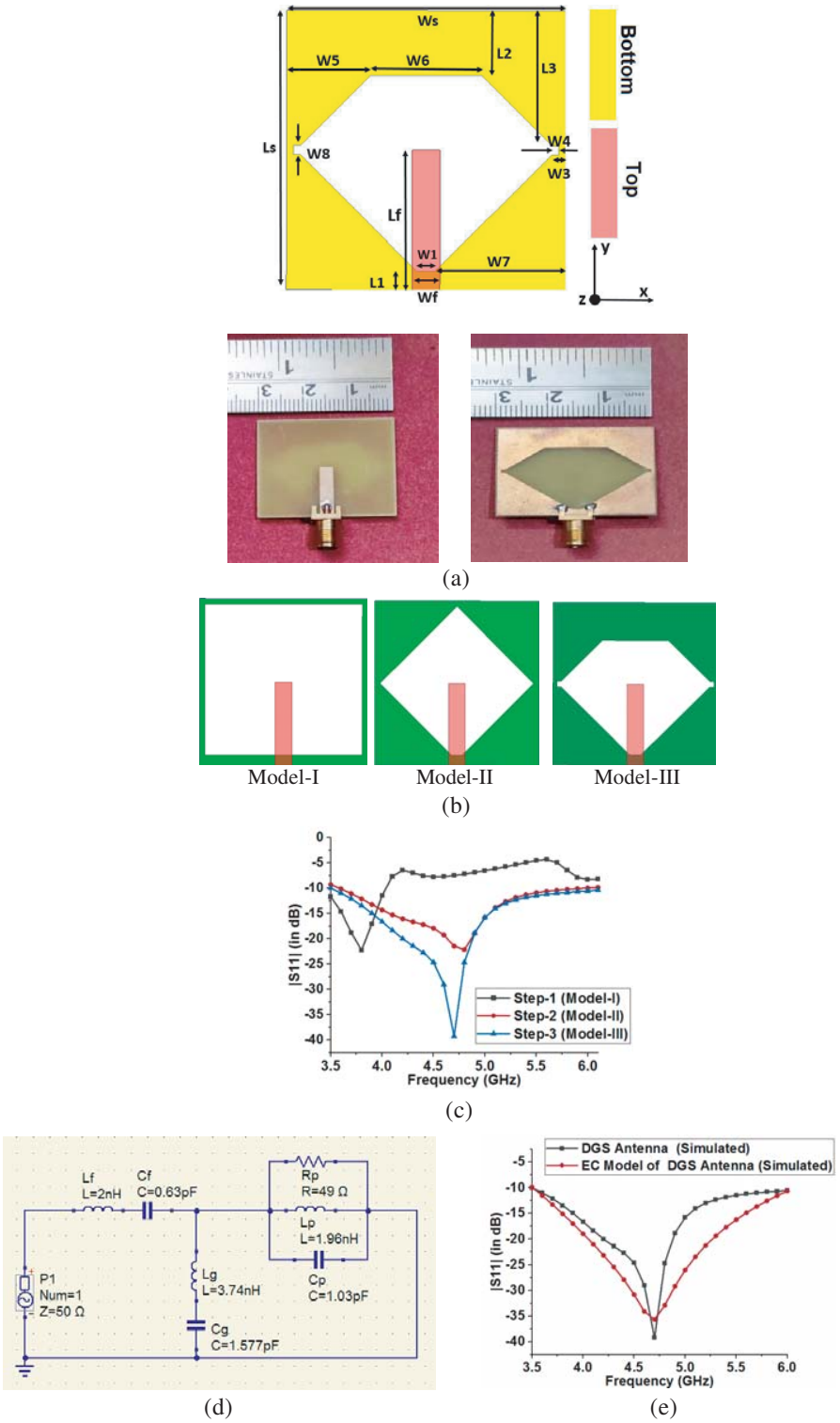
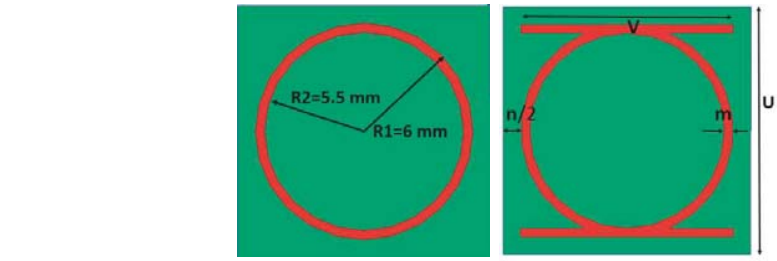
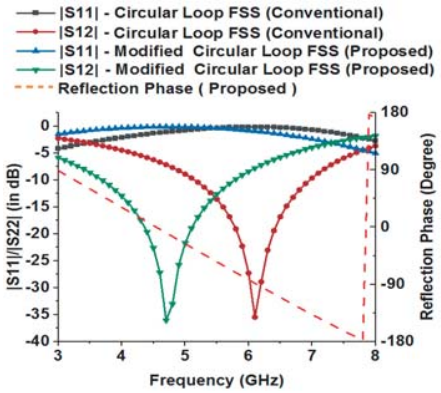


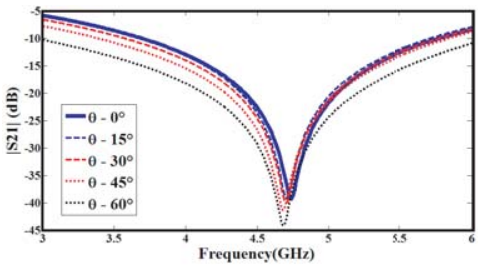
Figure 1. (a) Antenna prototype and fabricated antenna. (b) Antenna design steps. (c) S-parameters of antenna steps. (d) EC Model. (e) s-parameter of EC model and slot antenna.



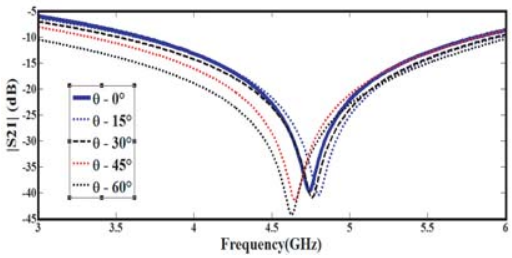
(a)



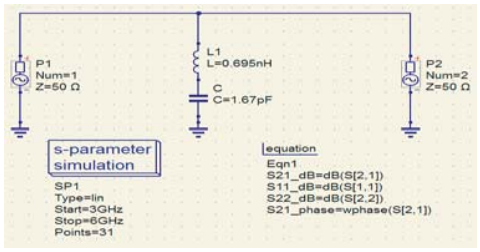
(b)



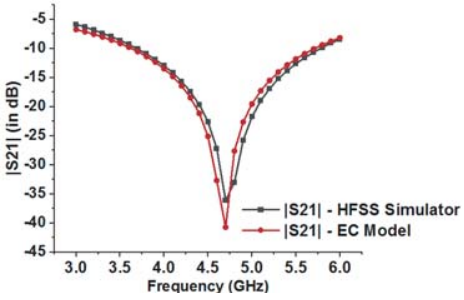
(c)



(d)



(e)



(f)

Figure 2. Proposed unit cell FSS. (a) Design steps. (b) s-parameters. (c) TE polarization. (d) TM polarization. (e) EC model. (f) s-parameters of EC model and with EM simulator.

[16]. The lumped elements of the EC model are extracted, and the optimized with QUCS software of the modified circular loop FSS are $L = 0.695\,\text{nH}$, $C = 1.67\,\text{pF}$ as shown in Fig. 2(e). The EC model and HFSS EM simulation are compared in Fig. 2(f), where the results validate the EC model with the

HFSS simulation.

$$B_c/Y_0 = wC = \frac{\pi}{2} 4\varepsilon_{eff} \frac{V}{U} F(U, n, \lambda, \theta) \tag{2}$$

$$X_L/z_0 = wL = \frac{\pi}{4} \frac{V}{U} F(U, 2m, \lambda, \theta) \tag{3}$$

where $F(U, n, \lambda, \theta)$ and $F(U, 2m, \lambda, \theta)$ can be computed from the

$$F(U, w, \lambda, \theta) = \frac{U}{\lambda} \cos\theta \left[\ln(\text{cosec}\frac{\pi w}{2U})\right] + G(U, w, \lambda, \theta) \tag{4}$$

$$G(U, w, \lambda, \theta) = \frac{(1-\delta^2)^2 \left[\left(1 - \frac{\delta^2}{4}\right)\left(\frac{A_+ + A_-}{2}\right) + 2\delta^2 A_+ A_-\right]}{\left(1 - \frac{\delta^2}{4}\right) + \delta^2 \left(1 + \frac{\delta^2}{2} - \frac{\delta^4}{8}\right)(A_+ + A_-) + 2\delta^6 A_+ A_-} \tag{5}$$

$$A_\pm = \left[1 \pm \frac{2U \sin\theta}{\lambda} - \left(\frac{U \sin\theta}{\lambda}\right)^2\right]^{-1/2} - 1 \tag{6}$$

and

$$\delta = \sin\left(\frac{\pi w}{2U}\right) \tag{7}$$

3.2. Performance Analysis of FSS over Microstrip Antenna

The FSS reflects the unwanted frequency coming from the antenna, and the maximum gain is achieved if the EM waves radiated from the antenna and the EM waves imitated from the FSS are in the same phase. The gain improvement of FSS loaded antenna is dependent upon the reflection phase ($\boldsymbol{\varphi_{FSS}}$) of FSS and height ($h$) of FSS superstrate from the antenna, extracted from Equation (8).

$$\boldsymbol{\varphi_{FSS}} - \mathbf{2}\boldsymbol{\beta}\mathbf{h} = 2\mathbf{N}\boldsymbol{\pi} \tag{8}$$

where $\mathbf{N} = 0, 1, 2, 3, \ldots$.

The extracted value of the height of FSS is 32.6 mm from the antenna by Equation (8). The gain enhancement characteristics have been achieved by 7×7 FSS unit cells with the same material, whereas one unit cell (Modified circular loop FSS) size is $14 \times 14 \times 1.6\,\text{mm}^3$. The FSS is loaded on a slot antenna with optimized distance as depicted in Fig. 3(a). $|S_{11}|$ in dB and a realized gain of the proposed FSS with a slot antenna are analyzed at different heights (h). The $|S_{11}|$ in dB and realized gain of an antenna with different FSS heights from 24.6 to 32.6 mm are presented in Fig. 3(b) and Fig. 3(c). Result reveals that when height of the FSS decreases, $|S_{11}|$ is shifted to the higher frequency due to an increase in coupling between FSS and antenna. The increasing height of FSS leads to decrease in the gain of the antenna at lower frequency and increased gain of the antenna at higher frequency. Higher gain enhancement of the antenna is achieved, when the height of the FSS from the antenna is reduced, but due to the coupling between the antenna and the FSS, the operating bandwidth of the antenna shifts towards higher frequency. Therefore, the optimized distance of the FSS from antenna is considered 28.6 mm. The 10 dB impedance bandwidth (IBW) at 28.6 mm height of FSS varies from 3.6 GHz to 6.1 GHz, which is close to operating frequency (3.5 GHz to 6.1 GHz) of the antenna without FSS. The maximum gain enhancement of 4 dB is achieved at 28.6 mm with less than 1.8 dB variation.

4. SIMULATED AND EXPERIMENTS RESULTS

The modified circular loop FSS with a slot antenna is designed on HFSS 13 and measured using Anritsu VNA (MS2038C). The $|S_{11}|$ (in dB) of the slot antenna with and without FSS are illustrated in Fig. 4(a). It is observed from the figure that the antenna exhibits frequency from 3.5 GHz to 6.1 GHz with an absence of FSS and 3.6 GHz to 6.1 GHz with the presence of FSS. The maximum realized gain is achieved as 7.87 dB and 4.3 dB of the proposed slot antenna with the presence and absence of FSS, as illustrated in Fig. 4(b), where the maximum 4 dB gain enhancement is achieved at 5 GHz. The gain

is reduced at 4.9 GHz due to disturbance of the surface current distribution near the feed and slot in the ground. Fig. 4(c) reveals that radiation efficiency (simulated) in the absence and presence of FSS is above 94% throughout the operating frequency band. The co- and cross-polar far-field normalized radiation patterns in xz and yz planes are measured in an anechoic chamber as illustrated in Fig. 4(d) and Fig. 4(e). The co- and cross-polarization difference is higher than 20 dB at resonant frequencies.

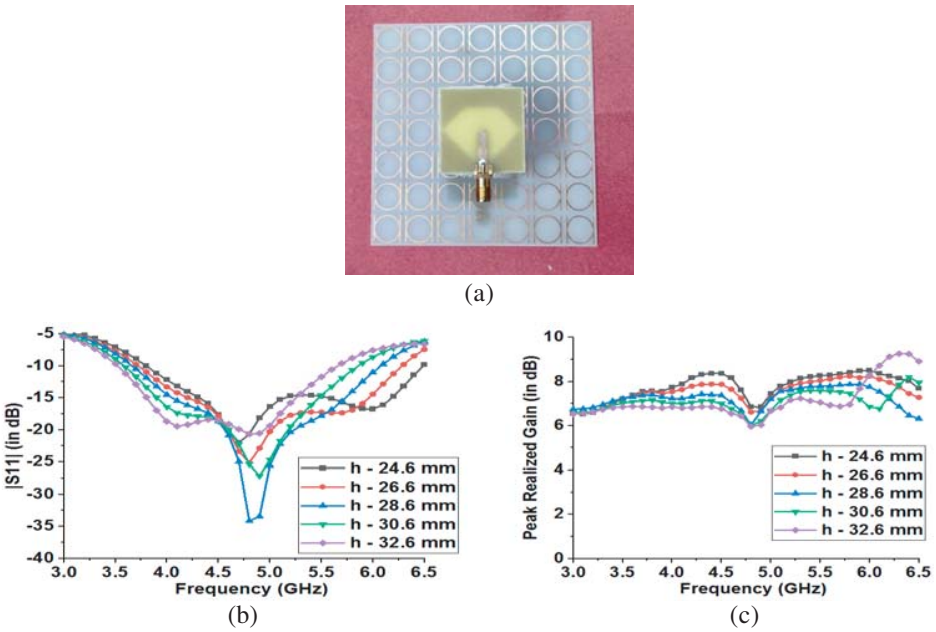
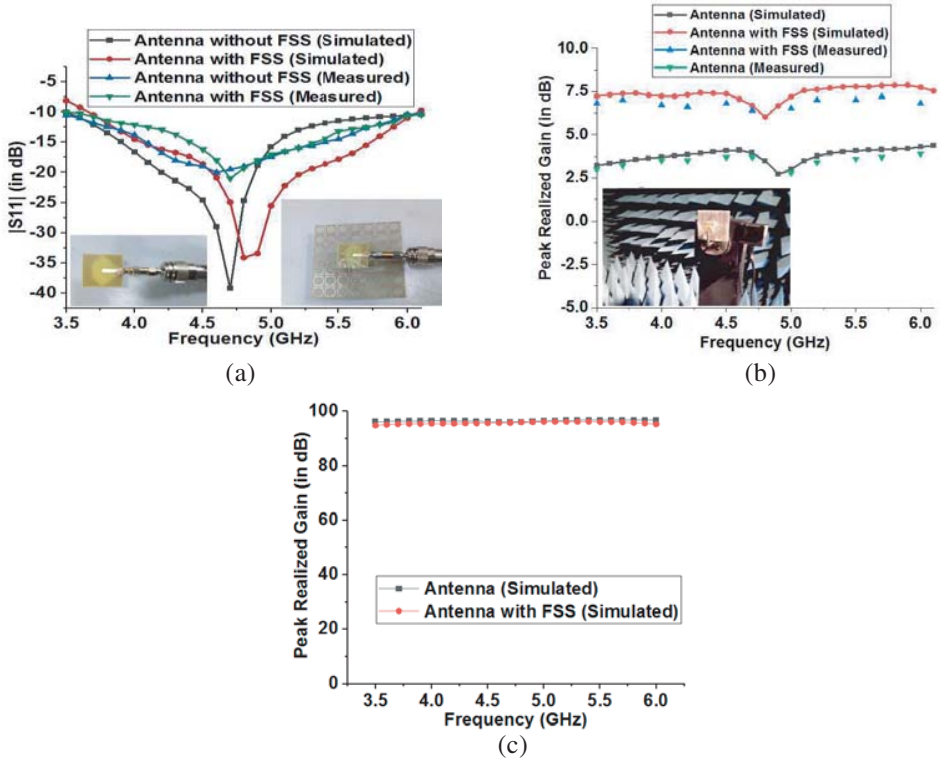


Figure 3. Slot antenna loaded at different heights of FSS layer. (a) Hardware prototype of FSS loaded on antenna. (b) s-parameters ($|S_{11}|$). (c) Realized gain (dB).



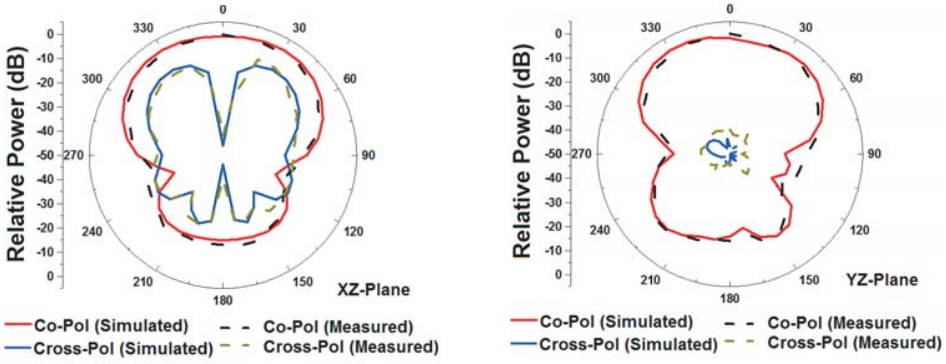


Figure 4. Slot antenna results with presence and without FSS. (a) S parameters ($|S_{11}|$ in dB, (b) Realized Gain (in dB). (c) Simulated radiation efficiency (%), Normalized radiation pattern in xz plane (d) and yz plane (e) at 4.7 GHz with FSS.

Table 1 describes the difference between the recently published papers, for gain enhancement. This table concludes that the proposed antenna manifests the 4 dB gain enhancement with a compact size of the antenna and single layer FSS. The maximum gain of the antenna with FSS is about 7.87 dB. Hence, the gain improvement of a slot antenna with a modified circular loop FSS is higher than those given by recently published articles [4, 6, 8–10] with single and dual-layer FSSs, where other articles [5, 7] used higher antenna size or double layer FSS.

Table 1. Gain enhancement comparison between proposed antenna and recently published FSS/metasurface papers.

Reference Nos.	Antenna Dimensions (mm^3)	Gain Enhancement Technology	Operating Frequency in GHz (Bandwidth)	Overall Gain	Improvement in gain (dB)
[4]	$70 \times 70 \times 1.6$	Band-Pass FSS	5–8 (3)	9	2–4
[5]	$60 \times 60 \times 1.6$	AMC	3.75, 5.98 and 8.79 (Multi-band)	7.52	3.88–4.93
[6]	$60 \times 60 \times 1.6$	Metamaterial	5.5–6.5 (1)	10	3
[7]	$63.65 \times 54.16 \times 1.58$	Band-Pass FSS (Double Layer)	2.4	7	5
[8]	$63.65 \times 54.16 \times 1.58$	SIW FSS	2.4	4.2	2
[9]	$50 \times 53.4 \times 1.6$	DRA-FSS	2.4	5.6	3
[10]	$40 \times 30 \times 0.8$ (DR height = 4.3 mm)	FSS (Double side)	8.15–13.2 (5.05)	7.8	3.3
[11]	$30 \times 60 \times 1.6$	FSS (Double Layer)	3.2–12 (8.8)	-	3–4
[12]	$63 \times 63 \times 0.8$	FSS (Double Layer)	3–10.6 (7)	9.8	4
Proposed	$30 \times 30 \times 1.6$	**Modified circular loop FSS (Single Layer)**	**3.6 to 6.1 (2.5)**	**7.87**	**4**

5. CONCLUSIONS

This paper presents a modified circular loop FSS loaded with a slot antenna for sub-6 GHz 5G applications. The modified FSS increases the stopband operating bandwidth and shifts it to a lower frequency band without the enhancement of FSS size. The modified FSS is insensitive to the phase change in TE and TM polarization modes. The slot antenna with FSS has obtained a bandwidth from 3.6 GHz to 6.1 GHz with up to 4 dB gain enhancement where the maximum gain of 7.87 dB is achieved.

REFERENCES

1. Sen, G., A. Banerjee, M. Kumar, and S. Das, "An ultra-wideband monopole antenna with a gain enhanced performance using a novel split-ring meta-surface reflector," *Microwave and Optical Technology Letters*, Vol. 59, No. 6, 1296–1300, 2017.
2. Kundu, S., A. Chatterjee, S. K. Jana, and S. K. Parui, "Gain enhancement of a printed leaf shaped UWB antenna using dual FSS layers and experimental study for ground coupling GPR applications," *Microwave and Optical Technology Letters*, Vol. 60, No. 6, 1417–1423, 2018.
3. Chatterjee, A. and S. K. Parui, "Frequency-dependent directive radiation of monopole-dielectric resonator antenna using a conformal frequency selective surface," *IEEE Transactions on Antennas and Propagation*, Vol. 65, No. 5, 2233–2239, 2017.
4. Chatterjee, A. and K. P. Susanta, "Gain enhancement of a wide slot antenna using a second-order bandpass frequency selective surface," *Radioengineering*, Vol. 24, No. 2, 455–461, 2015.
5. Ghosh, A., T. Mandal, and S. Das, "Design of triple band slot-patch antenna with improved gain using triple band artificial magnetic conductor," *Radioengineering*, Vol. 25, No. 3, 442–448, 2016.
6. Gharsallah, H., L. Osman, and L. Latrach, "Circularly polarized two-layer conical DRA based on metamaterial," *Microwave and Optical Technology Letters*, Vol. 59, No. 8, 1913–1919, 2017.
7. Belen, M. A., "Performance enhancement of a microstrip patch antenna using dual-layer frequency-selective surface for ISM band applications," *Microwave and Optical Technology Letters*, Vol. 60, No. 11, 2730–2734, 2018.
8. Güneş, F., M. A. Belen, and P. Mahouti, "Performance enhancement of a microstrip patch antenna using substrate integrated waveguide frequency selective surface for ISM band applications," *Microwave and Optical Technology Letters*, Vol. 60, No. 5, 1160–1164, 2018.
9. Belen, M. A., P. Mahouti, and M. Palandöken, "Design and realization of novel frequency selective surface loaded dielectric resonator antenna via 3D printing technology," *Microwave and Optical Technology Letters*, Vol. 62, No. 5, 2004–2013, 2020.
10. Bhattacharya, A., B. Dasgupta, and R. Jyoti, "Design and analysis of ultrathin X-band frequency selective surface structure for gain enhancement of hybrid antenna," *International Journal of RF and Microwave Computer-Aided Engineering*, e22505, 2020.
11. Krishna, R. R. and R. Kumar, "Slotted ground microstrip antenna with FSS reflector for high-gain horizontal polarisation," *Electronics Letters*, Vol. 51, No. 8, 599–600, 2015.
12. Ranga, Y., L. Matekovits, K. P. Esselle, and A. R. Weily, "Multioctave frequency selective surface reflector for ultrawideband antennas," *IEEE Antennas Wirel. Propag. Lett.*, Vol. 10, 219–222, 2011.
13. Roy, S. and U. Chakraborty, "Gain enhancement of a dual-band WLAN microstrip antenna loaded with diagonal pattern metamaterials," *IET Communications*, Vol. 12, No. 12 1448–1453, 2018.
14. Huang, J., T.-K. Wu, and S.-W. Lee, "Tri-band frequency selective surface with circular ring elements," *IEEE Transactions on Antennas and Propagation*, Vol. 42, No. 2, 166–175, 1994.
15. Langley, R. J. and E. A. Parker, "Equivalent circuit model for arrays of square loops," *Electronics Letters*, Vol. 18, No. 7, 294–296, 1982.
16. Varkani, A. R., Z. H. Firouzeh, and A. Z. Nezhad, "Equivalent circuit model for array of circular loop FSS structures at oblique angles of incidence," *IET Microwaves, Antennas & Propagation*, Vol. 12, No. 5, 749–755, 2017.
17. Das, P. and K. Mandal, "Modelling of ultra-wide stop-band frequency-selective surface to enhance the gain of a UWB antenna," *IET Microwaves, Antennas & Propagation*, Vol. 13, No. 3, 269–277, 2019.

Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021)
IEEE Xplore Part Number: CFP21K74-ART; ISBN: 978-0-7381-1327-2

Generating Caption for Image usingBeam Search and Analyzation with Unsupervised Image Captioning Algorithm

Prashant Giridhar Shambharkar
Department of Computer Engineering
Delhi Technological University
New Delhi, India
prashant.shambharkar@dtu.ac.in

Priyanka Kumari
Department of Computer Engineering
Delhi Technological University
New Delhi, India
priyankakumari_2k17co244@dtu.ac.in

Pratik Yadav
Department of Computer Engineering
Delhi Technological University
New Delhi, India
pratikyadav_2k17co236@dtu.ac.in

Rajat Kumar
Department of Computer Engineering
Delhi Technological University
New Delhi, India
rajatkumar_2k17co255@dtu.ac.in

2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) | 978-1-6654-1272-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICCS51141.2021.9432245

***Abstract*—In today's world of social media, almost everyone is a part of social platform and actively interacting with each other through internet. People on social media upload many pictures on their social media accounts with different captions. Thinking about the appropriate caption is a tedious process. Caption is important to effectively describe the content and meaning of a picture. Caption describes the image in meaningful sentences. A model for image caption generator can be built which is used to generate caption for images of different types and resolutions. Image captioning model which is used to generate caption in language that is understandable by a human being for the input images. CNN(convolution neural network) and RNN(recurrent neural network) is used using the concept of encoder-decoder to build this model. As CNN is used for image feature extraction purpose where only the important features or the important pixels, if the image is considered in the form of matrix of pixels, which are extracted from the resultant image, instead of CNN model, other pre-trained imagenet models which have higher accuracy will be used and their results are then compared by using BLEU score metric for comparison. For the prediction of captions, beamsearch method and argmax method is used and compared. The above discussed supervised image caption model is also comparedwith the built unsupervised image captioning model.**

The flickr8k dataset and then MSCOCO dataset are used to train and test the model. This model if implemented with the mobile application, which can be very useful for differently abled people, who completely rely on the assistance of text-to-speech feature.

***Index Terms*—Image caption generator, deep learning, CNN, RNN, LSTM, supervised and unsupervised learning, beam search, Argmax.**

I. INTRODUCTION

Image caption generator being a very famous and accepted field of research and study defines a way to build a model which can generate an English language based sentence or caption that is understandable by a human brain i.e., the sentence will be both syntactically and semantically correct. The resultant model will be able to describe the image precisely to a human brain.

In recent times, the studies and research related to captioning an image or providing the image with a meaningful sentence describing the image using some advanced deep learning techniques and algorithms and the concept of computer vision have become immensely popular and advanced.

Getting the machine ready to automatically describe the objects in the picture next to it or given on the display can be a difficult work to do, but its effect is immense in many field. As an example it can be used for the visually impaired people and can be helpful to them and can act as a guidance system for their day to day life. It does not only describe the image but also make user understand about it in an understandable language.

It is the very reason, the animation field is taken seriously in this generation. This paper has the scope of copying or mimicking the brain's capability to describe the image using natural meaningful sentence and analyze various process and due to this it is a very good problem for the field of artificial intelligence which included heavy deep learning including concept of computer vision and it has resulted in many organizations studying the concept of image caption generator.

Image captioning is often used for a wide range of cases used as a way to help blind people use text to speech with real-time output about the merging of the camera feed and the particular environment, to improve public awareness by converting photo captions into social feeds as speech mes-

Identify applicable funding agency here. If none, delete this.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:30:29 UTC from IEEE Xplore. Restrictions apply.

sages. Captions for each image on the web can create quick and precise image search and targeting. For robots, the agent's natural view is often given context by the display of natural language in the captions of images in the camera's center space.

This paper has used the encoder-decoder concept and methodology using CNN-RNN pair of architecture with beam search [1]. Here, RNN has used to provide input as the resultant optimized output of the CNN model instead of LSTM (Long Short Term Memory) to avoid the vanishing gradient problem. LSTM has the problem of vanishing gradient if the value of the parameter weight is taken to be less than 1, due to which the training of the model is improper and the resultant output is not as accurate as required. So, RNN model is used instead of LSTM where the value of weight is taken to be exactly equal to 1, if its value is greater than 1, the problem of overfitting may arise.

In this paper instead of building the CNN structure from the scratch for the optimization of the input image which is present as the matrix of pixels, various pre- trained models like Inceptionv3 and VGG16 are trained on heavy dataset resulting in a higher accuracy or lower validation loss/loss and compared their results for accuracy using BLEU (Bilingual Evaluation Understudy) scores. BLEU score is the metrics or standard to measure and identify how much understandable a sentence is to the human brain.

After covering the supervised methodology for image caption generator, we have compared the unsupervised manner for image caption generator for the accuracy of the results. The summary of our contribution to this paper:

- Built model for image caption generator with beam search using pre-trained model like VGG16 and Inceptionv3 that are already trained on huge datasets for higher accuracy.
- To resolve the "out of memory" issue, the size of batch during the training of the dataset is reduced.
- As the nature of the algorithm for image caption generator is stochastic, this may lead to slightly different results every time we run the model for results. So, random seed is set to generate same results everytime.
- We have used higher value of k in beam search for more efficient and accurate results with less validation loss/loss.
- In this paper, we compared the results of supervised and unsupervised methodology of image caption generator

II. Related Work

This particular section includes investigation of the part of the work that has recently been tried in this difficult area. Pre-image captioning techniques depend on the substitutionof a productive production model for text production in acommon language. Farhadi et al. (2010) [2], use trios next to a predetermined structure to produce text. They discriminate against the multi-name Markov Random Field to predict the measurement of trios. Kulkarni et al. (2011) [8], object is separated from the image, anticipate a multitude of features and related words (local data contradicting different articles) for each object, create a diagram called Basic Random Field and construct descriptive captions using marks and order. Furthermore, in the study, language parsing based model which are more efficient and powerful is also been studied and used [4, 5, 3, 6, 7]. These methods are inaccessible as they neglect to show a hidden subtitle regardless of whether each item was available in the configuration details. Likewise, the issue of format is their formal assessment.

For approaching and resolving this problem, the model is being implemented and used with the deep neural network, which has the vector space shared by both the image and the caption. Here, the Convolution Neural Network has been used as a typical approach[11] along with recurrent neural network approach [10] to the problem and it develop a number of high-lights that are used in sequence modeling process that studies the model for natural language to produce an understandable language showcase. We are different from their killings in order to improve the stable conditions. The Show, Attend and Tell uses new improvements in machine description and article recognition to introduce a thought-based model that looks at a few "spots" in a picture while making transcripts. They removed the highlights from the lower layer of convolution in contrast to the earlier layers, bringing the vector element $14 \times 14 \times 512$ in each image. We defined the number as the "196" spots in the picture, each with a vector of 512 objects. These 196 locations are included in the drafting process. Using visual considerations, the model had the option of finding out how to adjust its visibility to the key elements in the images when making subtitles. They introduced two instruments of thought, a "critical" object for processing prepared with back-to-back techniques; and a "solid" stochastic research program designed to maximize the hypothetical variability. Adding consideration improves the type of documentation produced, but at a higher cost of additional teachable loads. In addition, the preparation and exposure of the installed takes a ton of computer time, appropriately bringing this useless approach to the ever-present programs in the shopper's wireless devices. Karpathy and Li (2015) [12] introduced a method that uses accessible image data sets for displaying their sentences to gain expertise through multi-purpose connections among display information and descriptive language.

The planned model is defined by integrating Convolution Neural Networks into the display information areas, bidirectional Recurrent Neural Networks in the language, and the strategic goal of adapting these approaches. The design of the Multimodal Recurrent Neural Network, in these developed lines, uses expert arrangements to draw the required graphic text. The orderly model produces the best in the class bringing the research found into Flickr8K, Flickr30K and MSCOCO [13] datasets. As there are various improvements and enhancements in the field of the frameworks, Etienne et al (2016) [14] introduced another approach to improvement, small image go through independent sequence training, which is a modified form of reinforce [15]. Thus, the problem is resolved and help to develop their own model of inseparable test standards such as BLEU, ROUGE, etc., rather than being a coincidence.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:30:29 UTC from IEEE Xplore. Restrictions apply.

III. METHODOLOGY AND ARCHITECTURAL ANALYSIS

A. Datasets and Evaluation

This part of the paper helps in understanding the image captioning and open source database and the various method for it. The important parts or element for all the inventions that has be created by humans are feed up by the data, statistics and obviously the system power. These keys are important for each other. There is a fact that the data can make a model more powerful, reliable and more effective for its usage. The showing of the image on the display is somewhat familiar with description of the machine, and its testing strategy extends to the machine description to create its own unique test rules.

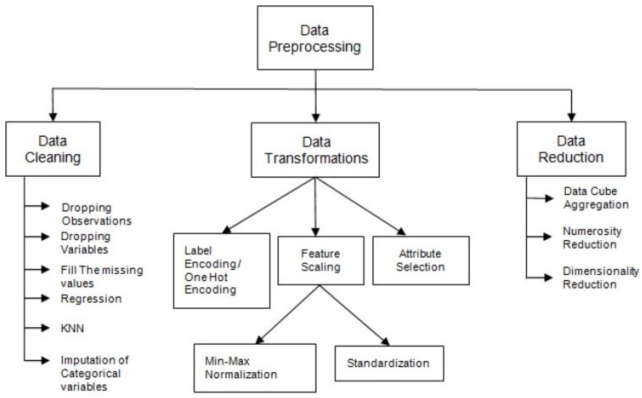
B. Data sets are accessible from making a caption generator

Information and knowledge are the basic keys the ideas going on in the human brain. It has been noticed that the rules and regulations that have a bulk of data and information in it are most likely to be unfollowed by the people. In earlier image caption generator methodology there were various datasets were used having large size or may be small size also, instance, Flickr8k, Flickr30k [17], MSCOCO, PASCAL, STAIR and AI Challenger Dataset. There are five images as reference in the database. To display the same image, there is a different grammar used so that the previous image can be displayed properly. For example, the Microsoft team created the Microsoft COCO Captions Database just to use the image captioning method to take generate the caption on daily basis and can be used to assign the tasks and the works, for example, classification and display, image recognition.

To build the required model of image captioning, we have used MSCOCO dataset having number of data/images equal to 82780.

C. Steps involved in data processing

Data purification: Raw data may miss data or have insignificant amounts of it. This can lead to inconsistencies in data. The data cleaning step handles all of these problems. All the values which are missing are rectified in accordance with the relevant terms and all non-essential amounts are deducted. Creating Test and Training Sets: This is one of the most important steps in data development. In this step we look at the data set in the training and setting of the test. For the training purpose a large portion of the data (70% -80%) is used and a small portion of the data (20% - 30%) is used for the experiment objective. Feature measurement: One variable dominates the other variable, if independent variables are measured at different scales. Feature measurement is therefore an important step in which all the values in the data are converted to a definite distance.



1. Flowchart

D. Model analysis

1) Supervised image captioning: For our paper, we will be using captions using CNN (Convolutional Neural Networks) and specialized RNN i.e. LSTM (Long Short- Term Memory). Image highlighting will be removed from inceptionv3/vgg16 model which is a CNN model based on image databases and thereafter feeds on the outstanding LSTM model which is responsible for producing the captions of a given image.

Convolutional Neural Network(CNN)
Convolutional Neural Networks are deep neural organizations that can deal with information in the form of input as a 2-D framework. We convert the images into a 2-D frame which are then fed into the CNN and then we can use it to successfully classify/identify the

given image. In this way, CNN provides us an easy way to work with images. We primarily use CNN for image orders, image recognition, flight or Superman, and so on. It filters images from left to right and rips to extract highlights from the image and merges part of the group photos. It can deal with translated, uplifted, enhanced images and changes in context.

Recurrent Neural Network(RNN)
Intermittent Neural Network is an assertion of a feed-forward neural organization with memory within. The RNN is intermittent as it plays the same power in each data offering while the yield of current information is based on a single previous calculation. Following the harvest, it is repeated and sent back to the repetitive organization. By resolving the options, it looks at current data and yields obtained from previous information.
Long-Term Memory (LSTM)
Long Short-Term Memory (LSTM) networks is a type of RNN, it is an altered version of duplicate neural networks, which is adjusted in a way so that it can remember past data in memory easily. The RNN disappearance issue is resolved. The LSTM is improvised upon RNN to categorize, process and improvise a series of time given a period of undetermined duration. Back distribution is used to train the model. In the LSTM network, there are three gates:

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:30:29 UTC from IEEE Xplore. Restrictions apply.

- Input gate: This gate helps determining the part of input data which is to be converted to memory. The function sigmoid determines which values you must allow to exceed 0,1. and the performance of tanh gives weight to the transferred values that determine their value level from 1 to 1.

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i) \quad (1)$$

$$C_t = tanh(W_c.[h_{t-1}, x_t] + b_C) \quad (2)$$

- Forget gate: Find out what information will be lost on the block. Determined by the function sigmoid. It observes the previous situation (ht-1)and input text(Xt)and subtracts a number in between0(leave this)and1 (keep this) for each of the numbers in styleCt-1.

$$f_t = \sigma(W_f.[h_{t-1}, x_t] + b_f) \quad (3)$$

- Output gate: The output of the LSTM structure is represented or provided by the this gate.

$$O_t = \sigma(W_o.[h_{t-1}, x_t] + b_o) \quad (4)$$

$$h_t = O_t * tanh(C_t) \quad (5)$$

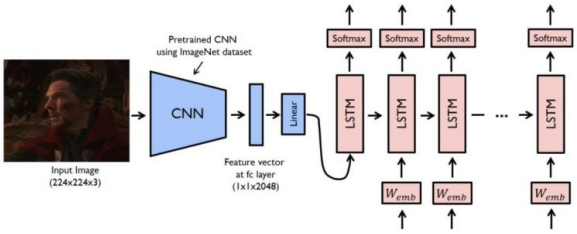


Figure 1. CNN-LSTM architecture for image captioning model

In this paper, the CNN model which is generally used for feature extraction of an image is replaced by VGG16/Inceptionv3 models as these models provide better accuracy and the image captions are predicted using beam search with different value of k and argmax search and then compared with each other for the accuracy and loss/validation loss. Flickr8k dataset and MSCOCO dataset is used here in this paper for experimentation.

- Approximately 1 hour is required for training of model/feature extraction for VGG16 model with Flickr8k dataset and 5-6 hours for MSCOCO dataset.
- Due to the fact that Inceptionv3 has very less parameter as compared to the VGG16 model so, it took 20 minutes for feature extraction with Flickr8k dataset and 2 hours 45 minutes for MSCOCO dataset using Inceptionv3 model.

2) Unsupervised image captioning: We have used the concept of machine translation using unsupervised technique for image captioning using unsupervised deep learning algorithm. We have considered out image as source language. We have used set of images I and sentences S^, and detector which detects visual concept. Ni is total numbers of images and Ns is total number of sentences. External corpus is used to get the sentence, they are unrelated to the image description.

The Model

The model is comprised of three parts , first one is image encoder, second is sentence generator and third is a discriminator for discriminating real sentences or produced by the generator.

Encoder

CNN model is used to provide the most important feature or pixels of an image as a result fim:

$$fim = CNN(I) \quad (6)$$

We used Inception-V4 for the purpose of encoder.

Generator

We have used LSTM for the purpose of generation of sentences. It takes the image representation and converts it into sentence in the language understandable by a human brain to describe the image. It provides us the probability distribution of all the words conditioned on the feature representation of image and the words that are already generated from our vocabulary.

The following equations are used to sample the generated word from the vocabulary.

$$x_{-1} = \mathrm{FC}(fim), \quad (7)$$
$$x_t = W_e s_t, t \in \{0 \ldots n-1\},$$
$$[p_{t+1}, h^g_{t+1}] = LSTM^g(x_t, h^g_t), t \in \{-1 \ldots n-1\},$$
$$s_t \sim p_t, t \in \{1 \ldots n\}$$

,

,

FC = Fully connected layer,

$\sim$ = operation used for sampling

n = generated sentence's length

W_e = word embedding matrix, x_t = input of LSTM

s_t = Representation of word generated by generator using one−hot

h^g_t = hidden state of LSTM

h^g_t = hidden state of LSTM

p_t = probability over the dictionary at the time step t

s_o = sentence starting point, s_n = sentence ending point

s_t is sampled from the probability distribution p_t

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:30:29 UTC from IEEE Xplore. Restrictions apply.

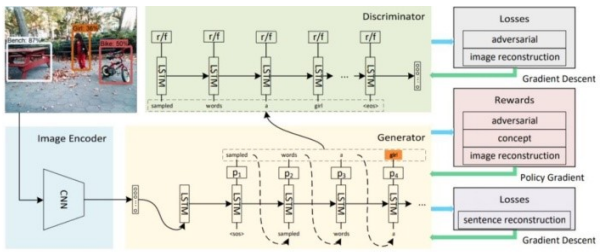


Figure 2. Model Architecture for unsupervised implementation of image caption generator.

Discriminator

We have used LSTM to implement the discriminator whose work is to discriminate the corpus contained real sentence from the sentence that the model generates.

$$[q_t, \square_t^d] = LSTM^d(x_t, \square_{t-1}^d), t \in \{1 \dots n\} \quad (8)$$

Training

As this is not supervised training. So, there are three objectives of the training in this paper.

- **Adversarial caption generation**
  An adversarial text generation method [22] is used in this paper. The image feature is passed as an input to the generator which generates a sentence best suited tothe feature of image.

  The main work of generator is to generate sentences as real as possible so that it can fool the discriminator. To accomplish this, a reward mechanism is used in each time-step which is known as "adversarial reward". The value for the reward for the t-th generated word can be calculated as:

$$r_t^{adv} = \log(q_t) \quad (9)$$

  Adversarial loss for the discriminator is defined as:

$$L_{adv} = -\left[\frac{1}{l}\sum_{t=1}^{l}\log(q_t^{\hat{}}) + \frac{1}{n}\sum_{t=1}^{n}\log(1-q_t)\right] \quad (10)$$

- **Visual Concept Distillation** The adversarial type of reward only ensures generation of plausible sentences but they may be irrelevant to the image. So it is important to identify all the visual concepts present in the image. We have used a visual concept detector for this purpose. When a word generated by the model where the image is used to identify and detect the analogous visual concept, a reward named as concept reward is given to the word which is generated, where the confidence score of the visual concept corresponds to the reward value. The visual concept detector takes an Image I and yields a set of different concepts and their confidence scores.

  C=(c1,v1),...,(ci,vi),...,(cnc,vnc)
  C= Confidence Scores
  ci= visual concepts corresponding to ith value
  vi= confidence score of ith detected visual
  N_c =total visual concepts
  The t-th generated word is assigned a concept reward which is given by:

- **Bi-directional Image-Sentence Reconstruction**
  Limited Number of object concepts can be detected reliably by the existing visual concept detectors. It's important for our model that it has a better understanding of semantic concepts of the image. The images and sentences are put together into a common latent space for reconstructing each other.

  Image Reconstruction
  We have used the sentence generated by the generator reconstruction of image feature instead of constructing the whole image. So as it can be seen in the figure 5 discriminator is also an encoder for sentences. A layer that is fully connected is arranged on the top of the discriminator so that the last hidden state can be papered to the latent space that is common to both sentences and images:

$$x' = FC(h_n^d) \quad (11)$$

  For the purpose of training the discriminator an additional image reconstruction loss is defined:

$$L_{im} = ||x_{-1} - x'||_2^2 \quad (12)$$

$$r_t^{im} = -L_{im} \quad (13)$$

  r_t^{im} = *Reward to the generator for reconstructing* the image

  L_{im} = *Reconstruction error*

  Sentence Reconstruction

  The derived representation can be used by the generator to reconstruct the sentences. It is also a type of sentence denoising auto-encoder [23]. The images and sentences are aligned by it in the latent space, and by having the common space for the representation of an image, the learning of the decoding of sentence takes place. The cross-entropy loss is defined as:

$$L_{sen} = -\sum_{t=1}^{l}\log(p(s_t = s_t^{\hat{}}|s_1^{\hat{}}, \dots, s_{t-1}^{\hat{}})) \quad (14)$$

  Integration

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:30:29 UTC from IEEE Xplore. Restrictions apply.

In this paper policy gradient [21] is used for training the generator, where the gradients are estimated with respect to all the trainable parameters. The gradients for the generator are provided by the sentence reconstruction loss using back-propagation. For updating the generator both of the gradients are used. The gradient is:

$$\nabla_\theta \mathscr{L}(\theta) = -\mathbb{E}\left[\sum_{t=1}^{n}\left(\sum_{s=t}^{n}\gamma^s\left(r_s^{adv} + \lambda_{c^{r_s^c}}\right) + \lambda_{im^{r_s^{im}}} - b_t\right)\right.$$

$$\left.\nabla_\theta \log(s_t^T p_t)\right] + \lambda_{sen}\nabla_\theta \mathscr{L}_{sen}(\theta) \quad (15)$$

The combination of image reconstruction and adversarial losses provide the updating required with the parameters with the help of gradient descent process within the discriminator.

$$L_D = L_{adv} + \lambda_{im} L_{im} \quad (16)$$

IV. Results and Discussion

In this paper, as pre-trained models has been used for the purpose of feature extraction which is then feed as the input vector to the RNN/LSTM model/architecture for the final image captioning. The model VGG16 has larger number of parameters as compared to the Inceptionv3 model that is why, the later model takes lesser time for the feature extraction/training of the input dataset.

For image caption prediction, two methods have been used in this paper, beam search and argmax search and the results are compared with respect to the different pre-trained models using tables. The criteria for the comparison is taken to be the loss/validation loss value instead of accuracy value and the standard metric for comparison used here is BLEU score.

Initialization A pipeline is proposed for the pre-training of the discriminator and the generator. For each of the training image we have generated a pseudo caption, then image captioning model is initialized by the pseudo image-caption pairs. A concept dictionary has been built which contains all the object classes present in the dataset OpenImages [20].

Here, sentence corpus is required and used to provide the required training of a concept to the sentence model. The concept of each image provided is effectively detected using existing visual concept detector. Then the training of generator has been done with the help of standardized deep learning algorithm which includes pseudo pairs of image caption[1].

Model and architecture	BEAM search method	Argmax method
VGG16 + AlternativeRNN Epochs = 18 Batch Size = 64 Optimizer = Adam	**k = 3** **BLEU Scores on Validation data** *(Higher the better)* BLEU-1: 0.593876 BLEU-2: 0.348569 BLEU-3: 0.242063 BLEU-4: 0.123221	**Crossentropy loss** *(Lower the better)* loss(train_loss): 2.2880 val_loss: 3.1889 **BLEU Scores on Validation data** *(Higher the better)* BLEU-1: 0.596655 BLEU-2: 0.342127 BLEU-3: 0.229676 BLEU-4: 0.108707
VGG16 + RNN Epochs = 7 Batch Size = 64 Optimizer = Adam	**k = 3** **BLEU Scores on Validation data** *(Higher the better)* BLEU-1: 0.568993 BLEU-2: 0.326569 BLEU-3: 0.226629 BLEU-4: 0.113102	**Crossentropy loss** *(Lower the better)* loss(train_loss): 2.6297 val_loss: 3.3486 **BLEU Scores on Validation data** *(Higher the better)* BLEU-1: 0.557626 BLEU-2: 0.317652 BLEU-3: 0.216636 BLEU-4: 0.105288

Table 1. BLEU scores based comparison for VGG16 model.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:30:29 UTC from IEEE Xplore. Restrictions apply.

Model and architecture	BEAM search method	Argmax method
InceptionV3 + AlternativeRNN Epochs = 20 Batch Size = 64 Optimizer = Adam	**k = 3** **BLEU Scores on Validation data** *(Higher the better)* BLEU-1: 0.606086 BLEU-2: 0.359171 BLEU-3: 0.249124 BLEU-4: 0.126599	**Crossentropy loss** *(Lower the better)* loss(train_loss): 2.4050 val_loss: 3.0527 **BLEU Scores on Validation data** *(Higher the better)* BLEU-1: 0.596818 BLEU-2: 0.356009 BLEU-3: 0.252489 BLEU-4: 0.129536
InceptionV3 + RNN Epochs = 11 Batch Size = 64 Optimizer = Adam	**k = 3** **BLEU Scores on Validation data** *(Higher the better)* BLEU-1: 0.605097 BLEU-2: 0.356094 BLEU-3: 0.251132 BLEU-4: 0.129900	**Crossentropy loss** *(Lower the better)* loss(train_loss): 2.5254 val_loss: 3.1769 **BLEU Scores on Validation data** *(Higher the better)* BLEU-1: 0.601791 BLEU-2: 0.344289 BLEU-3: 0.230025 BLEU-4: 0.108898



Figure 3. Comparing captions for images for different value of k for beam search method and argmax method used.

Here, in figure 3 we can observe that beam search method for prediction of caption of an image provide more accuracy and preciseness than the argmax method and within the beam search method, higher the value of k, better the results will be obtained. Above we discussed the result and discussion part of the supervised image captioning now in the later part we are going to discuss the outcome of the image captioning model when implemented through unsupervised manner.

- Shutterstock1 is used here to gather the image description which is eventually used to collect the corpus of sentence. Here, visual concept detector is used by using object detection model [27] which is properly trained on OpenImages [20].
- The model built for image caption generator using unsupervised algorithm provide enhanced and better results for accuracy with 28.9% if it is measured with respect to CIDEr score.
- The space dimension for shared latent and the dimensions which are available for LSTM are taken to be 512 which is a fixed value and the to achieve the approximately same level and scale of weighting parameters with respect to different rewards, the value of sen, im, c, are taken to be 1, 0.2,10 and 0.9 respectively.
- The training of the model or the dataset is done using Adam optimizer[26] and it has the learning rate of 10-4. In the initial phase, the optimization/minimization of the loss or we can say it a cross entropy loss is achieved by taking into consideration the learning rate to be 10-3. Consequently, during the testing of the model the beam search method is opted with the value of k taken to be 3.
- We have used both supervised learning and unsupervised learning to build the model of image caption generator and results were very noticeable and promising but as we have used sentence corpus not image sentence pairs in unsupervised learning so the captions are more diversified. Our vocabulary is huge in unsupervised leaning. Captions generated in unsupervised learning are more plausible and can be seen in figure 4.

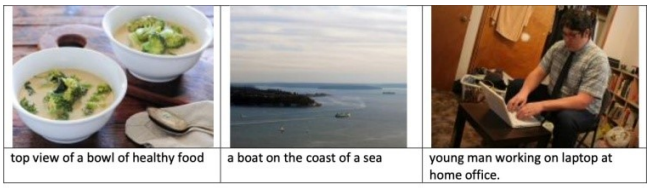


Figure 4. Qualitative image caption analysis for unsupervised manner for image caption generator.

V. CONCLUSION

As various earlier studies in the area/field of image captioning model has been done providing the state-of-art model which primarily generates captions for image having the architecture of encoder-decoder (i.e., CNN-RNN pair). Here, enhancement in accuracy and reduction in loss/validation loss

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:30:29 UTC from IEEE Xplore. Restrictions apply.

is achieved as some of the already trained imagenet models are taken into consideration for the feature extraction phase of the model building. The comparison among results for these models is done for the accuracy and then different methods used for the prediction of captions for input image are also compared. It is observed that the model is more improved with respect to the memory issue and accuracy if batch size of the dataset during the training of the model is taken to be of higher value and the value of k is higher in beam search method. The unsupervised method for image captioning is also experimented and model is built for it which shows improvement in the performance and accuracy of the model for image caption generator. As corpus which is an image describing content consist of over two million sentences gathered from Shutterstock used efficiently for the purpose of implementing unsupervised algorithm for image captioning that subsequently lead to the promising, more efficient and accurate results even if the image sentence label is not provided.

REFERENCES

[1] O. Vinyals, A. Toshev, S. Bengio and D. Erhan, "Show and tell: A neural image caption generator," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015, pp.3156-3164, doi: 10.1109/CVPR.2015.7298935.

[2] A. Farhadi, M. Hejrati, M. A. Sadeghi, P. Young, C. Rashtchian, J. Hockenmaier, and D. Forsyth. Every picture tells a story: Generating sentences from images. In ECCV, 2010.

[3] A. Aker and R. Gaizauskas. Generating image descriptions using dependency relational patterns. In ACL, 2010.

[4] M. Mitchell, X. Han, J. Dodge, A. Mensch, A. Goyal, A. C. Berg, K. Yamaguchi, T. L. Berg, K. Stratos, and H. D. III. Midge: Generating image descriptions from computer vision detections. In EACL, 2012.

[5] P. Kuznetsova, V. Ordonez, A. C. Berg, T. L. Berg, and Y. Choi. Collective generation of natural image descriptions. In ACL, 2012.

[6] P. Kuznetsova, V. Ordonez, T. Berg, and Y. Choi. Treetalk: Composition and compression of trees for image descriptions. ACL, 2(10), 2014.

[7] D. Elliott and F. Keller. Image description using visual dependency representations. In EMNLP, 2013.

[8] G. Kulkarni, V. Premraj, S. Dhar, S. Li, Y. Choi, A. C. Berg, and T. L. Berg. Baby talk: Understanding and generating simple image descriptions. In CVPR, 2011.

[9] R. Kiros and R. Z. R. Salakhutdinov. Multimodal neural language models. In NIPS Deep Learning Workshop, 2013.

[10] S. Hochreiter and J. Schmidhuber. Long short-term memory. Neural Computation, 9(8), 1997.

[11] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In arXiv:1502.03167, 2015.

[12] Andrej Karpathy, Li Fei Fei (2015) Deep Visual-Semantic Alignments for Generating Image Descriptions. IEEE Transactions on Pattern Analysis and Machine Intelligence (April 2017), vol 39, issue 4:664 –676

[13] T.-Y. Lin, et al (2014) Microsoft COCO: Common objects in context. arXiv:1405.0312

[14] Steven J. Rennie, Etienne Marcheret, Youssef Mroueh, Jarret Rossand Vaibhava Goel (2016) Self-critical Sequence Training for Image Captioning. in arXiv:1612.00563

[15] Ronald J. Williams (1992) Simple statistical gradient-following algorithms for connectionist reinforcement learning. In Machine Learning, pages 229–256, 1992.

[16] Marc' Aurelio Ranzato, Sumit Chopra, Michael Auli, and Wojciech Zaremba. (2015) Sequence level training with recurrent neural networks. ICLR, 2015.

[17] P. Mathur, A. Gill, A. Yadav, A. Mishra and N. K. Bansode, "Camera2Caption: A real-time image caption generator," 2017 International Conference on Computational Intelligence in Data Science(ICCIDS), Chennai, India, 2017, pp. 1-6, doi: 10.1109/ICCIDS.2017.8272660.

[18] Haoran Wang, Yue Zhang, Xiaosheng Yu, " An Overview of Image Caption Generation Methods", Computational Intelligence and Neuroscience, vol. 2020, Article ID 3062706, 13 pages, 2020. https://doi.org/10.1155/2020/3062706.

[19] T.-Y. Lin, et al (2014) Microsoft COCO: Common objects in context. arXiv:1405.0312.

[20] Ivan Krasin, Tom Duerig, Neil Alldrin, Vittorio Ferrari, Sami Abu- El-Haija, Alina Kuznetsova, Hassan Rom, Jasper Uijlings, Stefan Popov, Shahab Kamali, Matteo Malloci, Jordi Pont-Tuset, Andreas Veit, Serge Belongie, Victor Gomes, Abhinav Gupta, Chen Sun, Gal Chechik, David Cai, Zheyun Feng, Dhyanesh Narayanan, and Kevin Murphy. Openimages: A public dataset for large-scale multi- label and multi-class image classification. Dataset available from https://storage.googleapis.com/openimages/web/index.html, 2017.

[21] Richard S Sutton, David A McAllester, Satinder P Singh, and Yishay Mansour. Policy gradient methods for reinforcement learning with function approximation. In NIPS, 2000.

[22] William Fedus, Ian Goodfellow, and Andrew M Dai. Maskgan: Better text generation via filling in the . In ICLR, 2018.

[23] Pascal Vincent, Hugo Larochelle, Yoshua Bengio, and Pierre-Antoine Manzagol. Extracting and composing robust features with denoising autoencoders. In ICML, 2008.

[24] M. Hodosh, P. Young, and J. Hockenmaier. Framing image description as a ranking task: Data,models and evaluation metrics. JAIR, 47, 2013.

[25] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In arXiv:1502.03167, 2015.

[26] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.

[27] Jonathan Huang, Vivek Rathod, Chen Sun, Menglong Zhu, Anoop Korattikara, Alireza Fathi, Ian Fischer, Zbigniew Wojna, Yang Song, Sergio Guadarrama, et al. Speed/accuracy trade-offs for modern convolutional object detectors. In CVPR, 2017.

[28] Ranganathan, G. "Real Life Human Movement Realization in Multimodal Group Communication Using Depth Map Information and Machine Learning." Journal of Innovative Image Processing (JIIP) 2, no. 02 (2020): 93-101.

[29] Bindhu, V., and Villankurichi Saravanampatti PO. "Semi-Automated Segmentation Scheme for Computerized Axial Tomography Images of Esophageal Tumors." Journal of Innovative Image Processing (JIIP) 2, no. 02 (2020): 110-120.

[30] S. Han and H. Choi, "Domain-Specific Image Caption Generator with Semantic Ontology," *2020 IEEE International Conference on Big Data and Smart Computing (BigComp)*, Busan, Korea (South), 2020, pp. 526-530, doi: 10.1109/BigComp48618.2020.00-12.

[31] Vijayakumar, T., and R. Vinothkanna. "Retrieval of complex images using visual saliency guided cognitive classification." J Innov Image Process (JIIP) 2, no. 02 (2020): 102-109.

[32] N. K. Kumar, D. Vigneswari, A. Mohan, K. Laxman and J. Yuvaraj, "Detection and Recognition of Objects in Image Caption Generator System: A Deep Learning Approach," *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, Coimbatore, India, 2019, pp. 107-109, doi: 10.1109/ICACCS.2019.8728516.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:30:29 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021)
IEEE Xplore Part Number: CFP21K74-ART; ISBN: 978-0-7381-1327-2

Genre Classification using Character Networks

Rahul
Department of Software Engineering
Delhi Technological University
Delhi, India
rahul@dtu.ac.in

Ayush
Department of CSE
Delhi Technological University
Delhi, India
ayush_2k17co86@dtu.ac.in

Divya Agarwal
Department of CSE
Delhi Technological University
Delhi, India
divyaagarwal_2k17co112@dtu.ac.in

Devika Vijay
Department of CSE
Delhi Technological University
Delhi, India
devikavijay_2k17co107@dtu.ac.in

2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) | 978-1-6654-1272-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICCS51141.2021.9432303

***Abstract*—Literature has always had an undeniable impact on society and culture. It facilitates a nuanced understanding of the world around us and serves to pass on valuable knowledge to future generations; hence it is no surprise that the analysis of literature, and especially genre classification, is an essential field of study. However, the attempt to classify literary works into various categories has always been a persistent challenge for scholars. Assigning genres to fictional literature, which have a narrative style, is a particularly challenging task. Most of the existing work in this field have used statistical parameters in their classification algorithms, which neglect a narrative structure. To solve this problem, this research work has used Character Networks, graphs that explore interactions and relationships among various characters of a literary piece, making them singularly relevant to fiction. This paper proposes a novel method for genre classification that utilizes character networks derived from the text of books and compares Eigenvectors of Laplacian matrices of these networks directly against each other. This research work studies about different interpretations of the same character sets to eliminate noise and focus on character interactions. The proposed method can classify genres with 67.4% accuracy and an 80.5% F1 score.**

Keywords—Genre classification, character network, graph extraction, graph analysis, information retrieval

I. INTRODUCTION

Genre refers to a style in works of art, that has a particular set of characteristics. The attempt to classify literature or literary works on some set of parameters has a long history. Despite centuries of work, there is no one consensus on how to identify genres, what rules govern them and how best to identify various genres and so on. Notwithstanding, the difficulty and vagueness of the field, genres and the subsequent classification of literature into various genres is ubiquitous in our society. Whether it be the traditional brick-and-mortar library system or the new digital system, genre classification is prominently seen. Fig 1 shows some prominent examples of works of some genres.

If genre as a subject is so vague, then why is it so prominently used in our lives? Genres give readers a basic idea of what to expect from a particular book when they pick. When an individual picks up a novel labelled as a 'mystery' or a 'romance' or an 'adventure' novel, they are expecting some structure in the character interaction, writing style, story structure that they may have seen from reading previous novels under similar categories.

This is where the divide between fiction and non-fiction arrives. By its very nature, non-fiction literary works are easier to classify as they are usually about certain topics in particular. Works of fiction, however, exist to present a story to evoke certain emotions and tend to have certain overarching themes and styles. Fictional works can also often have elements in them that may lead to the works being classified under more than one category. For example, the 'Lord of the Rings' series can be categorized as both 'fantasy' and 'adventure' novels. Therefore, in fiction, there is a need to make the classifications as generalized as possible.

The analysis of literary work is a broad field involving many methods such as analysis of writing style, character interaction, vocabulary and so on. Using various machine learning and deep learning techniques, one can computationally analyse texts to find patterns that are not distinguishable by the human mind. This research work utilizes the character graphs to find out patterns in a story structure through the character interactions and using those patterns to identify the genre of the text.

A character network is a graph extracted from a novel, play [1] or movie [2] script that aims to represent the narrative structure of the work. Here, characters are represented by vertices and interactions between them are represented by the edges of the graph. This opens a new avenue to study relationships, interactions and the importance of various characters throughout the story. It is worth noting that such networks may be dynamic in nature [3] and this only adds to the multitude of information that can be derived from them.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:35:37 UTC from IEEE Xplore. Restrictions apply.

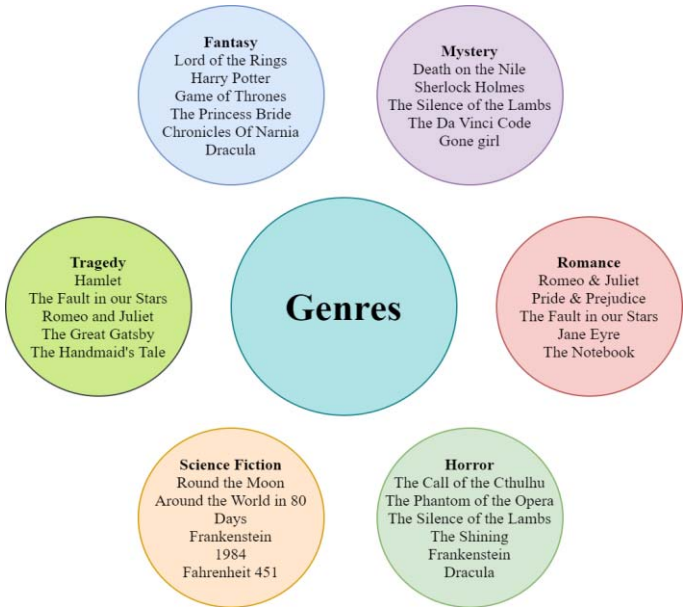


Figure 1 Genres and some major works of each genre

AI techniques are frequently utilized to help better understand and process education systems. Kumar, NM Saravana et al. [4] display such an implementation. Several narrative-related problems have been addressed through the analysis of character networks [5], such as

- Author style classification
- Genre classification
- Summarization
- Role detection.
- Information retrieval systems
- Recommendation systems
- Finding obscure patterns that can help in fields of humanities (e.g., sociology)

In this research, we are working with fanfiction as our input data. Fanfictions are literary works that are produced/written by fans of movies and books and are therefore related to the book's/ movie's 'universe' and have all/some characters from that universe. Our dataset has been pulled from Fanfiction.net [6], one of the world's largest fanfiction websites. Established in 1998, it has currently over 12 million registered users and contains stories in over 40 languages

II. Related Work

In the domain of computational research, an upcoming area is literature character network or graph analysis, which is essentially a social network containing numerous distinct features. A character network graph represents the interconnected social system created by the characters and leads us to explore complex social relationships and the intricacies of human behaviour displayed by the characters through their interactions.

Character Networks have been used to perform narrative analysis in many research fields. Many tasks which require a deep understanding of the narrative of the text have been handled using character graphs, including genre classification [7], authorship identification [8], lead character detection and analysis [9] [10], story/plot analysis [11] and community polarity detection [12].

T. Chakraborty and P. Choudhury [8] used graph-based methods to determine the authorship of Bengali literary works. They worked on a dataset consisting of the works of six Bengali authors. Authors used graph-based techniques over the traditional lexico-syntactic features, which are suitable only if there is a large dataset of documents. They used the interactions of characters, sentence structures and phraseological patterns as parameters to create graphs for each author. The paper showed that character networks could work better than traditional statistical methods in some applications

S. Muhuri et al. [12] applied weighted degree, closeness, and centrality of extracted character graphs to analyse the correlation among the characters of two renowned Bengali dramas of Rabindranath Tagore, Raktakarabi and Muktodhara. From the important nodes in the character graphs, features of the main characters, including the protagonist and antagonist, were analysed. Character clustering was also done with a high degree of relationship using a game theory-based community detection method. The paper highlighted that character networks could provide a mathematical relation for the interactions between characters

X. Zhao et al. [13] also utilized the technique of extracted character networks to identify the literary emotional polarity of characters and character communities in Chinese literature which helped clarify their underlying meaning. Due to the complex nature of character interactions and relationships, identifying polar opposite characters has always been a rather challenging task. From the polar vertices and polar edges of the graph, literary characters and communities were identified and divided. This method of community splitting based on graph polarity, capable of building the entire graph and inherently identifying positive and negative subgroups, was found to be much more accurate than the simple use of a co-occurrence graph. This work shows that character graphs could encapsulate complex character interactions.

L. Hettinger et al. [7] worked on the genre classification of German novels using character graphs and social interaction graphs extracted from the novels. They approached novel genre classification using three types of features - 1) features based on stylometrics, 2) features based on content, and 3) features based on social networks graphs and character interactions, the last being a novel set of features not previously applied to genre classification. They explored how these features affected genre classification performance of various machine learning classification algorithms and found that Support Vector Machines in combination with the topic-based features performed best. This was a fascinating result since genre and topic are often considered to be orthogonal.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:35:37 UTC from IEEE Xplore. Restrictions apply.

This method is based upon Character networks but does not offer a direct comparison between graphs that could be a more accurate representation of ground truth.

III Proposed Work

We propose a novel application of character networks to be used for genre classification for different interpretations of the same characters. Character networks provide a mathematical representation of character interaction data from a book. In previous work, researchers have tried to extract stylometric and statistical features from these graphs to predict the genre. We utilize the Eigenvectors of Laplacian Matrices of these graphs to compare them directly against each other to arrive at a more accurate representation of their content. We also eliminate the problem of shifting character sets by utilizing fanfictions. This provides us a unique opportunity to study the same character set under diverse interpretations.

A. Dataset

Fanfiction.net [6] is among the largest collection of stories written by fans available free of charge for anyone to read. The largest community on the site is Harry Potter with over 8,20,000 works of fiction spanning over 50 languages and 20 genres. n random books each of length 5,00,000 words or more were selected. All the books are processed to generate character network graphs in the following section.

The dataset was comprised of the following features

- Title of the book
- The complete text of the book
- Author of the book
- Genre of the book. Which is one of Humor, Family, Romance, Parody, Crime, Tragedy, Suspense, Supernatural, Spiritual, Angst, Hurt-Comfort, Horror, Sci-Fi, Fantasy, Adventure, Mystery, Friendship, Drama
- Language the book was written in.
- Date of publication of the book
- Number of chapters in the book

B. Character Network Extraction

Character Network Extraction refers to the conversion of raw text into its corresponding character network. The process involves NLP preprocessing that makes the text more suitable for machine consumption, identification of various characters, calculation of interaction of every character with every other character, conversion of character interaction data into a character graph and finally comparison of these graphs. Each of the steps is explained further in the next sections.

1) Preprocessing

A book is first processed through a tokenizer that splits the document by phrase, sentence, paragraph and each word. The resultant text is lemmatized using a language model that reduces each word to its base form, eliminating any inflection or derivative form of the word. This step reduces the complexity for further processing. The document is passed through a part of speech (POS) tagger [14] that assigns one of the parts of speech to a given word. This task of labelling each word allows for easy determination of proper nouns in the next step.

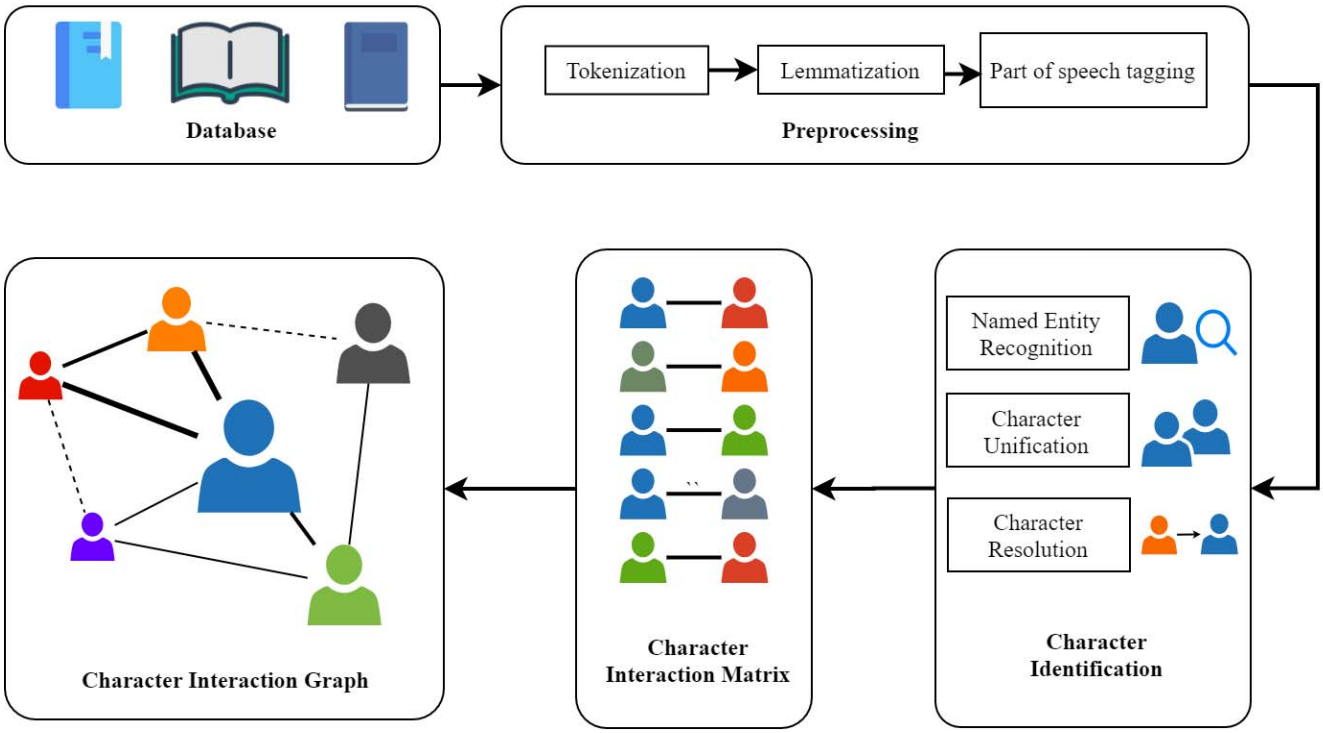


Figure 2 Generation of Character Interaction Graphs

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:35:37 UTC from IEEE Xplore. Restrictions apply.

2) Character Identification

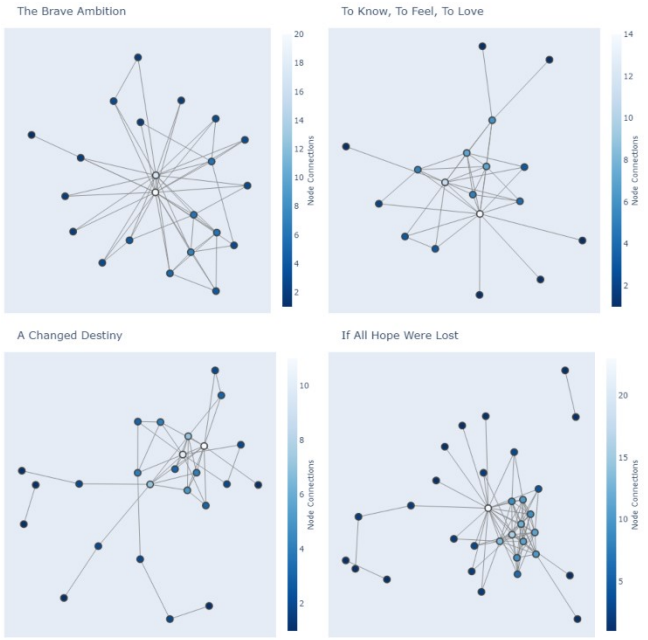


Figure 3 Graphs generated from stories in the dataset

Character Identification involves identifying characters in the text and linking their various occurrences throughout the story. These come in two major varieties: Proper nouns (like "Harry Potter") and Pronouns (like "He"). These are identified using POS tags obtained in the last step and are handled through Named Entity Recognition (NER) tools [15]. Various occurrences of the same characters also need to be linked up to ensure that if the character is referred to by different names, we still attribute to the original character (like "Harry Potter" can often be referred to as "Harry" or "Potter"). Once a list of characters is established and all occurrences are properly attributed, we can proceed to observe the interactions between the characters in the next step.

3) Character Interaction Matrix

The next step in the generation of character network graphs involves counting how often pairs of characters interact with one another. This is usually achieved by iterating over each sentence and marking which characters are present in that sentence. All variations of one character must be unified into a single character and marked accordingly. To obtain the matrix for the complete story, the sentential data is aggregated into a list of pairs of characters along with the count of interaction. A threshold is often applied to remove very small interactions and focus the graph on the narrative style. This list of pairs of characters and their interaction score can then be turned into a graph in the next step.

4) Graph Generation

Once the characters and their interactions have been established, the character interaction graph can now be generated using this data. We define vertices of the graphs to be characters and edges to be indicative of the presence of interaction among these characters. The edges are weighted by the interaction score calculated in the previous segment.

Fig 2 shows all the steps involved in the process of Character Interaction Graph Generation from books in the database. Fig 3 shows some of the graphs generated from the dataset. These graphs will be compared in the next step.

5) Graph Comparison

Graphs are not directly comparable on their own. Some might have more or fewer vertices, edges or weights. Hence we need a standard procedure that we can apply to a given graph to convert it into something directly comparable to other graphs. The spectrum of the Laplacian matrix of a graph is such a procedure.

For a graph G with n vertices, Laplacian matrix $L_{(n\,x\,n)}$ is defined in (1):

$$L_{i,j} = \begin{cases} deg(v_i) & if\ i = j \\ -w_{i,j} & if\ i \neq j \\ 0 & otherwise \end{cases} \quad (1)$$

where $deg(v_i)$ is the degree of vertex i and $w_{i,j}$ is the weight of edge between nodes i and j.

To compare two graphs G_1 and G_2, we calculate eigenvectors of laplacian matrices L_1 and L_2 constructed from G_1 and G_2. Closeness between G_1 and G_2 can be calculated as (2)

$$Closeness(G_1, G_2) = \sum_{i=1}^{min(len(\lambda_1), len(\lambda_2))} (\lambda_{1,i} - \lambda_{2,i})^2 \quad (2)$$

where the λ_1 eigenvector of G_1 and λ_2 is the eigenvector of G_2. Fig 4 shows the process of comparing two graphs and generating a closeness score that is further used in genre classification in the next step.

C. Genre Classification

All the graphs generated are then compared against each other using the difference in the sum of eigenvalues of the Laplacian matrix of the graphs as a measure of closeness. For each book, k closest graphs are selected from the entire dataset

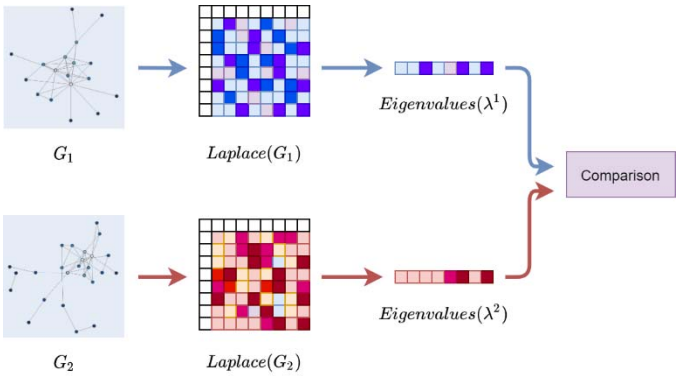


Figure 4 Comparison of two character graphs using laplacian eigen vectors

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:35:37 UTC from IEEE Xplore. Restrictions apply.

to serve as genre data and are arranged in non-decreasing order of closeness. The procedure is repeated for each book in the dataset resulting in an $(n_{samples} \times k)$ matrix that acts as input data for classifiers. Each classifier is run on this dataset wherein it tries to classify each sample based on its k closest samples.

1) Support vector machine (SVM)

Support vector machines try to find the hyperplane that best separates the given data points into their respective classes. For data that is not linearly separable, we modify the original data space through a dot product with a kernel function. The most common kernel function used is the Radial basis function kernel (RBF). In our tests, we have used both simple SVM (Linear SVM) and SVM with RBF kernel (RBF SVM).

The time complexity of Linear SVM is $O(n_{features})$ while that of RBF kernel SVM is between $O(n_{features} \cdot n^2{}_{samples})$ and $O(n_{features} \cdot n^3{}_{samples})$ based on the dataset.

2) Gaussian Process Classifier

Gaussian Process Classifier (GPC) utilizes exponential function and the gaussian distribution to classify input data into various classes. The aim is to estimate the parameters that could best differentiate among the given data.

The time complexity of Gaussian Process Classifier is $O(n^3{}_{samples})$

3) Multi-Layer Perceptron

Multi-Layer Perceptrons are modelled after neurons in the brain. Each Perceptron receives input from several perceptions in the previous layer. It performs a weighted sum of each of these inputs and uses a non-linear activation function to generate output for the next layer (Fig 5). All perceptrons of one layer are connected to each perceptron of the next layer, forming Neural Networks. The values of weights and biases of these perceptrons are random at initialization, but over subsequent epochs of training, the weights and biases are slowly updated using gradient descent through a method called backpropagation. MLP often outperforms most learning techniques but requires a large amount of data to be useful.

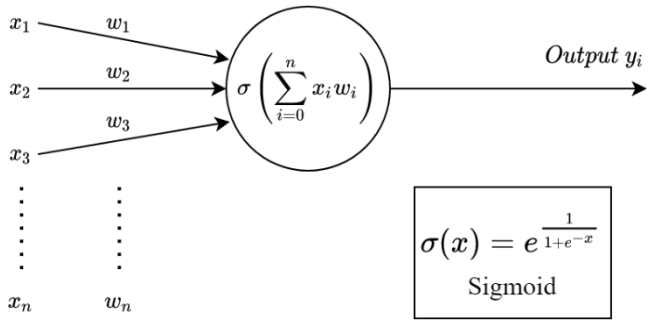


Figure 5 Single perceptron in MLP

The time complexity of MLP with single output neuron, h hidden layers, each with m neurons is

$$O(n_{sample} \cdot n_{features} \cdot m^h)$$

4) Random Forest

Random forest works under the assumption that a decision made by a collection of weak learners is better than a decision made by a single learner. Hence an ensemble of many shallow decision trees is created and each tree is trained on the data. Decisions are taken by taking each tree into equal consideration.

The time complexity for random forest classifier is $O(n_{samples} \cdot n_{features} \cdot log(n_{samples}))$

5) AdaBoost

Adaboost builds on the ideas of random forest further by introducing stumps. While individual decision trees in Random Forest still have multiple levels, decision trees in AdaBoost only have a single node and two leaf nodes. The importance of incorrect classifications is modified overtime to make the model better at learning tricky cases. This algorithm takes a decision as a weighted average of all the trees based on performance during training.

The time complexity of the AdaBoost classifier is

$$O(n_{samples} \cdot n_{features} \cdot n_{trees})$$

6) Gradient Boosting

Gradient Boosting Classifier optimizes for a differentiable loss function. It also works on the principle of meta estimator i.e., collection of weak individual learners who together form a much stronger learner. Here the trees have a fixed maximum size and misclassifications of previous trees are taken into consideration by subsequent trees.

The time complexity of Gradient Boosting Trees is $O(n_{samples} \cdot n_{features})$

7) Gaussian Naive Bayes

Naive Bayes Classifiers are based upon Bayes Theorem all while assuming that each feature is completely independent of any other feature and it makes an equal contribution to the outcome. Bayes theorem helps the classifier to learn and adjust its expectations when presented with new evidence. Gaussian Naive Bayes takes it one step further and assumes that each feature is distributed as per a normal distribution. The conditional probability thus becomes solvable by estimating the values of characteristic parameters of normal distribution, mean and std deviation in (3).

$$P(x_i | y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} exp\left(-\frac{(x_i - \mu_y)^2}{2\sigma_y^2}\right) \qquad (3)$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:35:37 UTC from IEEE Xplore. Restrictions apply.

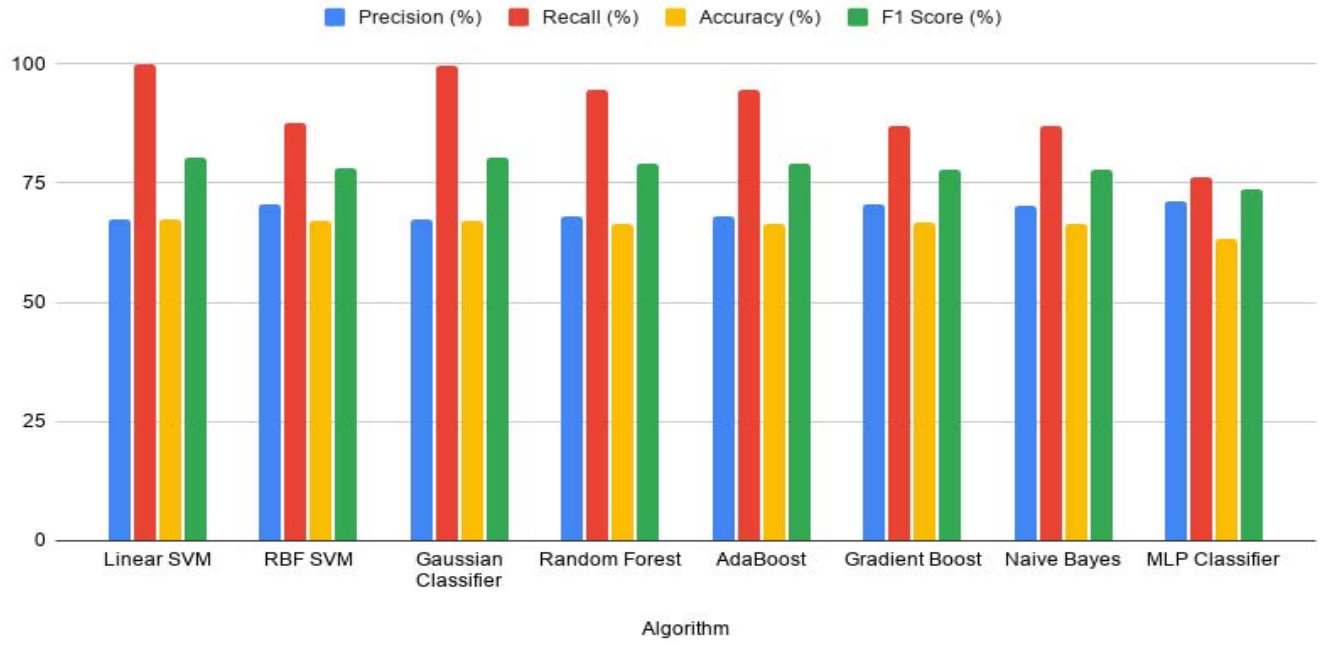


Figure 6 Chart representing results of genre classification

Where x_i is the event, y is the prior and σ_y and μ_y are standard deviation and mean of distribution of y respectively.

The time complexity of the Gaussian Naive Bayes algorithm is $O(n_{samples} \cdot n_{features})$

IV Results

In this work, a dataset of 500 books is used to classify genre using eight different classifiers. All the experiments were performed using python 3.6 on Intel i5 2.5GHz processor using scikit learn [16]. Dataset was split up into a training set (80%) and test set (20%). Each algorithm was run a total of five times achieving 5 fold cross-validation. Fig 3 shows some of the graphs generated from the dataset. In this representation, the length of the edge between two nodes is inversely proportional to the character interaction score, thus characters who have more interaction in the book are closer in the graph.

TABLE I. Genre classification using eigenvalues of laplacian matrix of character graphs as measure of closeness. (five fold cross validation, averaged)

Algorithm	*Precision (%)*	*Recall (%)*	*Accuracy (%)*	*F1 Score (%)*
Linear SVM	67.4	**100**	**67.4**	**80.52**
RBF SVM	70.54	87.51	67	78.13
Gaussian Classifier	67.33	99.7	67.2	80.37
Random Forest	68.13	94.63	66.6	79.21
AdaBoost	68	94.63	66.4	79.12
Gradient Boost	70.56	86.9	66.8	77.86
Naive Bayes	70.41	86.87	66.6	77.73
MLP Classifier	**71.17**	76.27	63.2	73.59

Precision, recall, accuracy and F1 scores were computed for each run and averaged out. Results have been tabulated in Table 1 and Fig 6. Support Vector Machine with Linear kernel comes out to be the model among the ones tested with an accuracy of 67.4% and F1 score 80.5%.

Genres are by nature very subjective and their classification and identification do not follow any unanimous rule or structure. What may be classified as science fiction, can also fall in the category of adventure. A single work of literature may have an overlap of different genres and the perceptions of the readers with regards to a single work may also differ. This makes categorising genres difficult and reduces the accuracy of the predictions. To give an example adventure and mystery may be perceived as a single genre. In the next section, we discuss some of the ways to improve upon the current model.

V Future Work

Pronouns are often very difficult to resolve in the context of a novel, especially when characters are conversing with dialogue written one after another. This problem is also present with the narrators in books written in a third-person perspective. The algorithm can be extended using Neural co referencing methods to better resolve these issues.

Recent advances in optimization techniques, network architecture and computation has allowed graph data to be processed using modern state of art neural networks. Character Interaction graphs generated in this paper can be analyzed using Graph Neural Networks to obtain more information about its characters, themes and plots.

Character Networks can also be used in conjunction with traditional feature-based methods for even better identification

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:35:37 UTC from IEEE Xplore. Restrictions apply.

of genre as well as the study of other narrative-related problems like role detection and summarization. [17] also explores other Neural Network Architectures that could be used along with the proposed method to provide a comprehensive depiction of fictional work.

VI CONCLUSION

In this paper, we have proposed the application of character networks to classify genres of books. We have used eigenvectors of the Laplacian matrices of the character graphs to compare graphs with the same character sets, thus emphasizing the character interactions. Character network analysis is emerging as a valuable tool in the toolbelt of data scientists analyzing the narrative structure of works of fiction. Genre classification is a prime candidate for the use of a character network as it can extract the narrative details and stylistic subtleties often missed by purely statistical methods. Character networks serve as an excellent foundation for recommendation and classification systems.

REFERENCES

[1] C. Suen, L. Kuenzel, and S. Gil. 2013. Extraction and analysis of character interaction networks from plays and movies. In Proceedings of the Digital Humanities. http://snap.stanford.edu/class/cs224w-2011/proj/laneyk_Finalwriteup_v1.pdf.

[2] Q. D. Tran, D. Hwang, and J. J. Jung. 2015. Movie summarization using characters network analysis. In Proceedings of the 7th International Conference on Computational Collective Intelligence. Lecture Notes in Computer Science, Vol. 9329. Springer, 390–399. https://doi.org/10.1007/978-3-319-24069-5_37

[3] A. Grener, M. Luczak-Roesch, E. Fenton, and T. Goldfinch. 2017. Towards a Computational Literary Science: A Computational Approach to Dickens' Dynamic Character Networks. Technical Report. Zenodo. https://doi.org/10.5281/zenodo.259499

[4] Kumar, NM Saravana. "Implementation of artificial intelligence in imparting education and evaluating student performance." Journal of Artificial Intelligence 1, no. 01 (2019): 1-9.

[5] Vincent Labatut and Xavier Bost. 2019. Extraction and Analysis of Fictional Character Networks: A Survey. ACM Comput. Surv. 52, 5, Article 89 (October 2019), 40 pages. DOI:https://doi.org/10.1145/3344548

[6] Fanfiction.net http://www.fanfiction.net

[7] L. Hettinger, M. Becker, I. Reger, F. Jannidis and A. Hotho, "Genre Classification on German Novels," 2015 26th International Workshop on Database and Expert Systems Applications (DEXA), Valencia, 2015, pp. 249-253, doi: 10.1109/DEXA.2015.62

[8] T. Chakraborty and P. Choudhury, "Authorship identification in Bengali language: A graph based approach," 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), San Francisco, CA, 2016, pp. 443-446, doi: 10.1109/ASONAM.2016.7752272.

[9] V. H. Masías, P. Baldwin, S. Laengle, A. Vargas, and F. A. Crespo. 2017. Exploring the prominence of Romeo and Juliet's characters using weighted centrality measures. Digit. Scholarsh. Hum. 32, 4 (2017), 837–858. https://doi.org/10.1093/llc/fqw029

[10] F. Jannidis, I. Reger, M. Krug, L. Weimer, L. Macharowsky, and F. Puppe. 2016. Comparison of methods for the identification of main characters in German novels. In Proceedings of the DH. 578–582. http://dh2016.adho.org/abstracts/297.

[11] O.-J. Lee and J. J. Jung. 2019. Modeling affective character network for story analytics. Future Generation Computer Systems 92 (2019), 458–478. https://doi.org/10.1016/j.future.2018.01.030

[12] S. Muhuri, S. Chakraborty and S. N. Chakraborty, "Extracting Social Network and Character Categorization From Bengali Literature," in IEEE Transactions on Computational Social Systems, vol. 5, no. 2, pp. 371-381, June 2018, doi: 10.1109/TCSS.2018.2798699

[13] X. Zhao, X. Liang, F. Tang, X. Zhang, X. Zhou and X. Guo, "Building Character Graphs and Dividing Communities in Chinese Novels Based on Graph Data Extraction: Community Division for Character Emotional Polarity Networks," in IEEE Access, vol. 8, pp. 95559-95573, 2020, doi: 10.1109/ACCESS.2020.2995738.

[14] Honnibal, M., & Montani, I. (2017). spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing.

[15] Srivastava, S., Chaturvedi, S., & Mitchell, T. (2016). Inferring interpersonal relations in narrative summaries. In 30th AAAI Conference on Artificial Intelligence, AAAI 2016 (pp. 2807–2813). AAAI press.

[16] Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

[17] Smys, S., Joy Iong Zong Chen, and Subarna Shakya. "Survey on Neural Network Architectures with Deep Learning." Journal of Soft Computing Paradigm (JSCP) 2, no. 03 (2020): 186-194

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:35:37 UTC from IEEE Xplore. Restrictions apply.

International Journal Of Multidisciplinary Research In Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 4.988

| Volume 4, Issue 5, May 2021 |

Green Synthesis of Nanoparticles and its Application in Sewage Treatment

Mikhlesh Kumari[1], Sunita Punia[2], Rajesh Kumar Lohchab[3] , Anil Kumar Haritash[4]

Department of Environmental Science and Engineering, Guru Jambheshwar University of Science & Technology, Hisar, Haryana, India[1,2,3]

Department of Environmental Engineering, Delhi Technological University, Delhi, India[4]

Supervisor[3], Co-supervisor[4]

ABSTRACT: Green synthesis of nanoparticles (NPs) is an emerging research trend in green nanotechnology as this method is nontoxic or less toxic, eco-friendly, efficient, and cost-effective as compared to other conventional physical and chemical methods. Green synthesis of NPs employs various biological agents such as plants, bacteria, algae, and fungi, but nowadays plant-based green synthesis of NPs is gaining more attention among researchers from around the world. A variety of green synthesized NPs are currently being used in water and wastewater treatment due to their high efficiency and biocompatible nature. Green synthesized NPs are highly proficient for recycling and removal of heavy metal from wastewaters without loss of their stability and degradation of a variety of organic pollutants from wastewaters and, thus, purify the wastewaters for reuse and recycling and could solve various water quality issues worldwide. However, regeneration and reusability are the main challenges to researchers and scientist yet in the green synthesis approach as a technology transfer from laboratory scale to commercial applications. Adsorption of organic pollutants, toxic metal ions, and removal of harmful bacteria can give us clean and pure drinkable water from wastewater resources. Respective magnetite nanoparticles (MNPs) were synthesized using a cheaper and greener way in an open-air environment with the use of crude latex of *Jatropha curcas* (JC) and leaf extract of *Cinnamomum tamala* (CT). Characterization of MNPs had been performed by dynamic light scattering (DLS), Ultraviolet-visible (UV-vis) spectroscopy, Fourier-transform infrared (FTIR) spectroscopy, powdered X-ray diffraction (XRD), and field emission scanning electron microscope (FE-SEM). The size ranges of the synthesized MNPs were observed in between 20–42 nm for JC-Fe_3O_4 and within 26–35 nm for CT-Fe_3O_4 by FE-SEM images. The effect of synthesized magnetic nanoparticles in wastewater treatment (bacterial portion), dye adsorption, toxic metal removal as well as antibacterial, antioxidant, and cytotoxic activities were studied. This purification will lead to an increase in the resources of pure drinking water in the future. In materials science, “green” synthesis has gained extensive attention as a reliable, sustainable, and eco-friendly protocol for synthesizing a wide range of materials/nanomaterials including metal/metal oxides nanomaterials, hybrid materials, and bioinspired materials. As such, green synthesis is regarded as an important tool to reduce the destructive effects associated with the traditional methods of synthesis for nanoparticles commonly utilized in laboratory and industry. In this review, we summarized the fundamental processes and mechanisms of “green” synthesis approaches, especially for metal and metal oxide [e.g., gold (Au), silver (Ag), copper oxide (CuO), and zinc oxide (ZnO)] nanoparticles using natural extracts. Importantly, we explored the role of biological components, essential phytochemicals (e.g., flavonoids, alkaloids, terpenoids, amides, and aldehydes) as reducing agents and solvent systems. The stability/toxicity of nanoparticles and the associated surface engineering techniques for achieving biocompatibility are also discussed. Finally, we covered applications of such synthesized products to environmental remediation in terms of antimicrobial activity, catalytic activity, removal of pollutants dyes, and heavy metal ion sensing. Due to the exceptional characteristics which resulted from nanoscale size, such as improved catalysis and adsorption properties as well as high reactivity, nanomaterials have been the subject of active research and development worldwide in recent years. Numerous studies have shown that nanomaterials can effectively remove various pollutants in water and thus have been successfully applied in water and wastewater treatment. In this paper, the most extensively studied nanomaterials, zero-valent metal nanoparticles (Ag, Fe, and Zn), metal oxide nanoparticles (TiO_2, ZnO, and iron oxides), carbon nanotubes (CNTs), and nanocomposites are discussed and highlighted in detail. Besides, future aspects of nanomaterials in water and sewage treatment are discussed.

I. INTRODUCTION

Green nanotechnology refers to the use of nanotechnology to enhance the environmental sustainability of processes producing negative externalities. It also refers to the use of the products of nanotechnology to enhance sustainability. It includes making green nano-products and using nano-products in support of sustainability.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 4.988

| Volume 4, Issue 5, May 2021 |

Green nanotechnology has been described as the development of clean technologies, "to minimize potential environmental and human health risks associated with the manufacture and use of nanotechnology products, and to encourage replacement of existing products with new nano-products that are more environmentally friendly throughout their lifecycle

Nanotechnology offers the potential of novel nanomaterials for the treatment of surface water, groundwater, wastewater, and other environmental materials contaminated by toxic metal ions, organic and inorganic solutes, and microorganisms. Due to their unique activity toward recalcitrant contaminants, many nanomaterials are under active research and development for use in the treatment of water and contaminated sites. Nanoremediation is the use of nanoparticles for environmental remediation[1] Nanoremediation has been most widely used for groundwater treatment, with additional extensive research in wastewater treatment. In 2017 the RingwooditE Co Ltd was formed in order to explore Thermonuclear Trap Technology (TTT) for the purpose of cleaning all sources of water from pollution and toxic contents. This patented nanotechnology uses a high pressure and temperature chamber to separate isotopes that should by nature not be in drinking water to pure drinking water, as to the by the WHO´s established classification. This method has been developed by among others, by professor Vladimir Afanasiew, at the Moscow Nuclear Institution. This technology is targeted to clean Sea, river, lake and landfill waste waters. It even removes radioactive isotopes from the sea water, after Nuclear Power Stations catastrophes and cooling water plant towers. By this technology pharmaca rests are being removed as well as narcotics and tranquilizers. Bottom layers and sides at lake and rivers can be returned, after being cleaned. Machinery used for this purpose are much similar to those of deep sea mining. Removed waste items are being sorted by the process, and can be re used as raw material for other industrial production. Perpetual exposure to heavy metal pollution and particulate matter will lead to health concerns such as lung cancer, heart conditions, and even motor neuron diseases. However, humanity's ability to shield themselves from these health problems can be improved by accurate and swift nanocontact-sensors able to detect pollutants at the atomic level. These nanocontact sensors do not require much energy to detect metal ions or radioactive elements. Additionally, they can be made in automatic mode so that they can be readably used at any given moment. Additionally, these nanocontact sensors are energy and cost effective since they are composed with conventional microelectronic manufacturing equipment using electrochemical techniques.[2]

Some examples of nano-based monitoring include:

1. Functionalized nanoparticles able to form anionic oxidants bonding thereby allowing the detection of carcinogenic substances at very low concentrations.[32]
2. Polymer nanospheres have been developed to measure organic contaminates in very low concentrations
3. "Peptide nanoelectrodes have been employed based on the concept of thermocouple. In a 'nano-distance separation gap, a peptide molecule is placed to form a molecular junction. When a specific metal ion is bound to the gap; the electrical current will result conductance in a unique value. Hence the metal ion will be easily detected."[35]
4. Composite electrodes, a mixture of nanotubes and copper, have been created to detect substances such as organophosphorus pesticides, carbohydrates and other woods pathogenic substances in low concentrations.

Although green nanotechnology poses many advantages over traditional methods, there is still much debate about the concerns brought about by nanotechnology. For example, since the nanoparticles are small enough to be absorbed into skin and/or inhaled, countries are mandating that additional research revolving around the impact of nanotechnology on organisms be heavily studied. In fact, the field of eco-nanotoxicology was founded solely to study the effect of nanotechnology on earth and all of its organisms. At the moment, scientists are unsure of what will happen when nanoparticles seep into soil and water, but organizations, such as NanoImpactNet, have set out to study these effects.[3]

II. DISCUSSION

Nanomaterials are very small-sized particles that have nanoscale dimension approximately from 1 to 100 nm. The green nanomaterials or nanoparticles are synthesized by biological methods. During the synthesis of green nanoparticles, natural materials such as plants, microorganisms and organic polymers like carbohydrates, proteins and fats are actively involved. Due to the very small size, nanomaterials are playing a significant role in wastewater treatment system and give up possibilities for targeting very specific points. Different types of biosynthesized nanomaterials are used in the controlling of water pollution, such as zinc oxide nanoparticles, phytogenic magnetic nanoparticles, biopolymer-coated metal nanoparticles, silver-impregnated cyclodextrin nanocomposites, silver nanoparticles in *Aloe vera* plant extract and *Photocatalytic titania* nanoparticles. Green nanoparticles are providing an

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 4.988

| Volume 4, Issue 5, May 2021 |

alternative way for elimination of pollutants from the water bodies. The application of green nanoparticles provides an economical, convenient and eco-friendly way of sewage treatment. With the increase in the global human population, the adequate supply of resources has become limited. The development of pollution-free technologies for environmental remediation and clean energy supplies for the sustainable growth of human society is the need of the hour. Nanotechnology can have a substantial impact on developing 'cleaner' and 'greener' technologies with significant health and environmental benefits. The applications of nanotechnology are being explored for their potential to provide solutions to manage, mitigate, and clean-up air, water, and land pollution, as well as to improve the performance of conventional technologies used in environmental clean-up. Green nanotechnology is the branch of nanotechnology that envisages sustainability through various applications.[4,5]

The present article deals with the topics related to green nanotechnology for sustainable development. The applications of nanotechnology used to solve environmental issues by reducing the overall energy consumption during the synthesis and manufacturing process, the ability to recycle products after use, and to develop and use eco-friendly materials were summarized in this aticle. The sections have been divided according to the applications of green nanotechnology. The nano-manufacturing processes, green synthesis of nanomaterials, and the treatment of wastewater with reference to the principles of green chemistry have been discussed in detail. Currently, nanotechnology shows great promise to solve the sustainability issues, but it is impossible to overlook the adverse effects of nanomaterials on the environment and human health. In spite of the high performance and low cost of nano-remediation technology, advanced research is necessary to understand and prevent the potential adverse environmental impacts, i.e., ecosystem-wide impacts. The present chapter highlights the green chemistry principles influencing the life cycle of nano-products from design to disposal. The various applications and limitations of green nanotechnology have been discussed in the light of green chemistry principles for sustainability.[6]

III. RESULTS

Sewage Treatment Methodologies

Nanophotocatalysts

The word "photocatalysis" is composed of two Greek words "photo" and "catalysis" which mean decomposition of compounds in the presence of light. Usually, in scientific world there is no consensus definition for photocatalysis [**7**]. However, this term can be employed to define a process to activate or stimulate the substance by using light (UV/Visible/Sunlight). Photocatalyst which changes the rate of reaction without any involvement by itself during the chemical transformation process. Furthermore, the key difference between traditional thermal catalyst and photocatalyst is that the prior is activated through heat while the final is activated through photons of light energy [**8**]. Nanophotocatalysts are commonly used for wastewater purification, as they help to enhance the reactivity of catalyst due to having a greater surface ratio and shape dependent features [**9**].

The nano size-based materials show different response as compared to bulk materials due to their distinct quantum effects and surface properties. It assists to increase their electric, mechanical, magnetic chemical reactivity and optical properties too [**8**]. It has been showed that the nanophotocatalysts can expand the oxidation ability due to effective production of oxidizing species at surface of material which helps in degradation of pollutants from the polluted water effectively [**7**].

Nano- and Micromotors

Nanotechnology, an area of research that has progressive at such a quick pace in early decades and offered many approaches for water treatment. In modern era, nano/micromotors have been considered that can convert the energy from several resources into machine-driven force, empowering them to achieve special goals through various mechanisms. These innovative motors are motorized in both cases by using fuel or without fuel sources (acoustics, magnetic field, electric field) have several significant exciting applications [**9**]. They show more speed, high power, specific control movement, self-mix ability, etc. The removal of contaminant from polluted resources is a significant focus for environmental stability and sustainability [**5**]. The trend of water purification and its treatment has grown quickly throughout the world because of high demands for pure water resources and novel water superiority regulations. A large variety of approaches are used earlier for removal of pollutants from polluted groundwater, fresh water, sediments wastewater, etc.

Nano/micromotors can go in nano/microscale detentions in the presence of a magnetic field, serving as programmed cleaners, where conventional approaches are not actively working. However, most positive concept for nano/micromotors in term of wastewater decontamination depends on fuels. There is problem, this condition reduced

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 4.988

| Volume 4, Issue 5, May 2021 |

the working potential of nanomotors in biological applications. Although some photocatalytic, biocatalytic and magnetically driven nano/micromotors have been industrialized, still there are some challenges in order to apply nano/micromotors for seawge treatment in future.

Nanomembranes

Nanomembranes are a unique type of membrane formed with different nanofibers which are employed to eliminate unwanted nanoparticles present in aqueous phase. Using this technique, the process takes places at a very high elimination speed with condensed fouling propensity and it also served as a pre-treatment process which is used for reverse osmosis [**11**]. There are many reported studies on membrane nanotechnology in order to produce multifunction membrane by using different nanomaterial substance in different polymers-based membranes. Water porous membrane for water treatment did nanofiltration, ultrafiltration, reverse osmosis, etc. The membrane contains a porous support with composite layer. Typically, the considerable composite layer is carbon-based material (graphene oxide/CNT) dispersed into polymer matrix for significant practice. This provide promising and significant progresses in fouling resistance and aqueous transport. For example, CNTs hold anti-microbial properties that can minimize fouling, biofilm formation and it can also reduce the chance of mechanical failures [**13**]. The doping process of nanomaterial (zeolite, alumina, TiO_2, etc.) into polymer ultra-filtration membranes show the formation of amplified membrane on surface of hydrophilicity and fouling resistance .The antimicrobial material like silver metal particles are also doped with a polymer to produce polymeric membrane to prevent attachment of bacteria and inhibit biofilm production on surface of membrane . It served to inactivate viruses and it has also ability to prevent the biofouling of membrane .The production and growth of nanomaterial thin film are incorporated into active thin film composite through doping in surface modification. Usually, there are a few major challenges in nanomembrane i.e., membrane clogging and membrane fouling. Therefore, it is a need to overcome this problem. By the addition of super hydrophilic nanoparticles in making a thin film nanocomposite membrane, the prevention from clogging and fouling could be checked. Furthermore, metal oxide nanoparticles (Al_2O_3, TiO_2), antimicrobial nanoparticles (nano silver, CNTs) and aquaporin-based membranes are very useful material to overcome the membrane clogging and fouling issue due to having high hydrophilicity, high porosity, better fouling resistance and a better homogenous nanopore. Generally, nanoparticles affect the selectivity and permeability of a membrane which depends upon the sort, quantity, dimension, etc. of nanoparticles. Moreover, there are many biological membranes present with highly permeable and selective ability . Nanomembranes are also used for wastewater treatment due to having several properties that make this material more prolific, these are high uniformity, homogeneity ability, optimization, short time required, easily handled and contain much order of reaction There are some nano photocatalysts which can be introduced in the nanocomposite membrane to make fit it for the degradation of organic pollutants. For example, TiO_2 incorporated nanomembranes and films are effectively used to deactivate different microorganisms and degrade the organic pollutants The developments and progressive growth of nanotechnology especially in nanomaterial produced several nanostructured catalytic membranes which contain novel properties like improved selectivity, high decomposition rate, and larger foul resistance .In order to synthesize these types of nanomaterials, there were many approaches used to produce it with multi-functionality features .The incorporation of nanostructured catalytic materials like iron-catalysed based free radical and enzymatic catalysis within pore membrane showed a constructive progress in this technology. Therefore, it possible to carry out oxidative reactions for removal of pollutants and detoxification of water without use of any toxic chemicals. To prove the efficiency of nanostructured catalytic membrane in industrial/commercial applications, immobilized membrane nanoparticle (ferrihydrite/iron oxide) reaction was carried out with hydrogen peroxide to make free radicals' ions for removal of chlorinated based organic pollutant in real groundwater. The development of nanostructured materials is still very useful in many other environmental applications [10] There were several reported studies of metallic nanoparticles immobilization on membrane (e.g., chitosan, cellulose acetate, polysulfones, etc.) for dichlorination, degradation of a toxic substance which contains novel properties such as, hindrance of nanoparticles, high reactivity, absence of agglomeration and surface reduction [11] Palladium acetate and polyetherimide both are used to prepare nanocomposite films and a novel type interaction is present among hydrogen and Pd nanoparticles in order to improve water treatment efficiency In situ and ex-situ methods are also used to produce nanomaterials within the matrix with precursor film under many conditions This offers many opportunities to produce nanomaterial with tunable properties. Moreover, the developments of nanotechnology are also responsible to produce many significant catalytic membranes with high selectivity, better permeability and high resistance fouling. Modern methods contain hybrid and bottom-up approaches for empowering its multi-functionality characteristics in the field of wastewater treatment [12]

Nanosorbents

Nanosorbents hold wide properties like high sorption capability that make the nanosorbents more appropriate and powerful for water treatment These nanosorbents are very rare in commercial form but researcher and experts doing a lot of work on it to produce nanosorbents in larger quantity/at commercial level The most commonly reported

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 4.988

| Volume 4, Issue 5, May 2021 |

nanosorbents are based on carbon (e.g., carbon black, graphite, graphene oxide). Furthermore, metal/metal oxide and polymeric nanosorbents were also exist The composite of different material like Ag/polyaniline, Ag/carbon, C/TiO_2, etc. carrying huge significant importance in order to reduce the effect of toxicity in the wastewater treatment process. The carboneous material such as CNTs with a cylindrical form nano structure may present as single-walled and multiwalled nanotubes depending on the method of synthesizes. CNTs hold measurable adsorption sites and due to high surface area, they hold sustainable surfaces. It must be stabilized to prevent from aggregation that decreases the surface-active sites because CNTs has hydrophobic surface properties. So, it is an efficient material for the pollutants by adsorbing process. Similarly, the polymeric nanoadsorbents like dendrimers are functional for eliminating organic pollutants and heavy metals from wastewater . For example, copper ions were reduced with the help of dendrimer-ultrafiltration system [13]. They simply regenerated by shifting of pH effect and show biocompatibility, biodegradability, and toxic free environment. Furthermore, the removal percent of dyes or others organic pollutants is almost 99% [10]. Another important nanosorbent is zeolites, which have an absorbent structure in which several nanoparticles like copper, silver ions can be implantable [9]

IV. CONCLUSION

Water distinguishes and makes our planet superior as compared to other planets. Though the worldwide available supply of pure water is high to meet all existing and predictable water demands. There are several areas where the drinking water resources are insufficient to fulfil the basic, economic and domestic developmental needs. In such areas, the insufficient fresh water to fulfil human water need and sanitation requirements is certainly a limit on human health and for other living creatures. Scientific community/research institutes must find a path to eliminate these limitations. Moreover, the world is facing several challenges in doing that, especially given a fluctuating and undefined future environment, a fast-rising population that is driving enlarged community and financial growth, urbanization and globalization. How superlatively overcomes on these challenges which entails exploration in all features of water management. The trend of nanomaterial for water pollutant treatment is rapidly increasing in this modern era due to very horrible conditions of water and demand for fresh water in the whole world. There is an important need for innovative progressive water treatment approaches, in specific to certify a high class of water for drinking purpose, remove micro/macro pollutants and increase industrial production developments through flexibly modifiable water treatment approaches. Nanotechnology has proved great achievement for controlling water purification challenges and makes some future advancement. Nanomaterial approaches like nanosorbents, nanostructured catalytic membranes, etc. are very efficient, less time required, less energy and eco-friendly techniques but all these methods are not cheap, and they are not used yet for commercial purpose to purify the wastewater at a large scale.

Nanomaterials show high efficiency due to having high rate of reaction. However, there are still some weaknesses that must be negotiated. Up to now, no operational digital monitoring techniques exist that offer consistent real-time measurement facts on the superiority of nanoparticles which are existing in small amounts in H_2O [**12**]. Furthermore, to reduce the health risk, some research institutes and international research communities should prepare proper guidelines to overcome this issue. Another, further mechanical restriction of nano-engineered water approach is that they are infrequently flexible to mass developments, and at present-day, in several cases are not modest with conservative treatment approaches [**13**]. However, nano-engineered materials provide excessive potential for water revolutions, in specific for decentralized water treatment technologies, point-of-use strategies, and seriously degradable pollutants. Furthermore, there is a great need to synthesize some modified nanomaterials which should be effective, having high efficiency, easy to handle and eco-friendly. It is also necessary to take the cost challenges and commercialization of these technologies for wastewater treatment. The different applications of nanomaterial can provide a tremendous offer in order to supply drinking water to whole world.

REFERENCES

1. Ahmad, A.; Mohd-Setapar, S.H.; Chuong, C.S.; Khatoon, A.; Wani, W.A.; Kumar, R.; Rafatullah, M. Recent advances in new generation dye removal technologies: Novel search for approaches to reprocess wastewater. *RSC Adv.* **2015**, *5*, 30801–30818. [**Google Scholar**] [**CrossRef**]
2. Zhang, Y.; Wu, B.; Xu, H.; Liu, H.; Wang, M.; He, Y.; Pan, B. Nanomaterials-Enabled water and wastewater treatment. *Nano Impact* **2016**, *3*, 22–39. [**Google Scholar**] [**CrossRef**]
3. Gitis, V.; Hankins, N. Water treatment chemicals: Trends and challenges. *J. Water Process Eng.* **2018**, *25*, 34–38. [**Google Scholar**] [**CrossRef**]
4. Hodges, B.C.; Cates, E.L.; Kim, J. Challenges and prospects of advanced oxidation water treatment processes using catalytic nanomaterial. *Nat. Nanotechnol.* **2018**, *13*, 642–650. [**Google Scholar**] [**CrossRef**] [**PubMed**]

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 4.988

| Volume 4, Issue 5, May 2021 |

5. Umar, K.; Haque, M.M.; Mir, N.A.; Muneer, M. Titanium dioxide-Mediated photocatalyzed mineralization of Two Selected organic pollutants in aqueous suspensions. *J. Adv. Oxid. Technol.* **2013**, *16*, 252–260. [**Google Scholar**]
6. Umar, K.; Ibrahim, M.N.M.; Ahmad, A.; Rafatullah, M. Synthesis of Mn-Doped TiO_2 by novel route and photocatalytic mineralization/intermediate studies of organic pollutants. *Res. Chem. Intermediat.* **2019**, *45*, 2927–2945. [**Google Scholar**] [**CrossRef**]
7. Baruah, S.; Khan, M.N.; Dutta, J. Perspectives and applications of nanotechnology in water treatment. *Environ. Chem. Lett.* **2016**, *14*, 1–14. [**Google Scholar**] [**CrossRef**]
8. Wu, Y.; Pang, H.; Liu, Y.; Wang, X.; Yu, S.; Fu, D.; Chen, J.; Wang, X. Environmental remediation of heavy metal ions by novel-Nanomaterials: A review. *Environ. Pollut.* **2019**, *246*, 608–620. [**Google Scholar**] [**CrossRef**]
9. Daer, S.; Kharraz, J.; Giwa, A.; Hasan, S.W. Recent applications of nanomaterials in water desalination: A critical review and future opportunities. *Desalination* **2015**, *367*, 37–48. [**Google Scholar**] [**CrossRef**]
10. Yaqoob, A.A.; Ibrahim, M.N.M. A Review Article of Nanoparticles; Synthetic Approaches and Wastewater Treatment Methods. *Int. Res. J. Eng. Technol.* **2019**, *6*, 1–7. [**Google Scholar**]
11. Tang, W.W.; Zeng, G.M.; Gong, J.L. Impact of humic/fulvic acid on the removal of heavy metals from aqueous solutions using nanomaterials: A review. *Sci. Total Environ.* **2014**, *468*, 1014–1027. [**Google Scholar**] [**CrossRef**]
12. Mir, N.A.; Haque, M.M.; Khan, A.; Umar, K.; Muneer, M.; Vijayalakshmi, S. Semiconductor mediated photocatalysed reaction of two selected organic compounds in aqueous suspensions of Titanium dioxide. *J. Adv. Oxid. Technol.* **2012**, *15*, 380–391. [**Google Scholar**] [**CrossRef**]
13. Umar, K. Water Contamination by Organic-Pollutants: TiO_2 Photocatalysis. In *Modern Age Environmental Problem and Remediation*; Oves, M., Khan, M.Z., Ismail, I.M.I., Eds.; Springer Nature: Basel, Switzerland, 2018; pp. 95–109. [**Google Scholar**]

Corrected Proof

Check for updates

Water Science & Technology

IWA PUBLISHING

© 2021 The Authors

Water Science & Technology Vol 00 No 0, 1 doi: 10.2166/wst.2021.241

Heat and water flux modeling in an earth dam

Sushant Kumar*, A. K. Sahu and Munendra Kumar

Department of Civil Engineering, Delhi Technological University, Delhi, India
*Corresponding author. E-mail: sushant_phd2k17@dtu.ac.in

ABSTRACT

This study aims to identify the water flux in an earth dam using heat flux due to convection. Sixteen earth dam models were constructed in a hydraulic flume by varying geometrical and flow input parameters to identify heat and water flux. Homogeneous as well as earth dams with clay cores were built in a hydraulic flume. Temperature measurements were done to calculate heat flux in the experimental model. A finite element model of the earth dam using Seep/w was developed to obtain water flux, while temp/w was used to obtain heat flux. These results were used as input in Temp/w and Seep/w in Geostudio 2020. Significant reduction of the heat and water flux was seen while comparing the homogeneous models with central impervious core models. An increase in the heat and water flux was observed on increasing the downstream filter's length, longitudinal slope, and vice versa with the upstream slope and the thickness of the clay core. Comparing fluxes in a homogeneous dam model (model 1) with the clay core model (model 9) with top width 2.4 m and bottom width 18 m in model 9, both water flux and heat flux were reduced by 78.46%. While comparing it with model 10, with bottom core width of 18 m and top core width of 1.9 m, both water flux and heat flux reduced by 77.72%. Heat flux measurements were found to be a valuable alternative to detecting water flux and seepage in an earth dam at a reduced cost.

Key words: convection, earth dam, heat flux, porous medium, temperature, water flux

HIGHLIGHTS

- This study identifies the heat due to the convective heat transfer in the respective earth dam model and water flux due to seepage in a porous media such as earth dams by varying geometrical and flow input parameters.
- A finite element model using Seep/w was used to determine the water flux, and Temp/w was used to obtain heat flux.
- It was observed that near walls of a hydraulic flume, soil pores of the earth dam were of larger volume, resulting in an increased local water flow.
- Increased heat and water fluxes were observed when the downstream filter length and longitudinal slope were increased, and vice versa with the upstream slope and thickness of the clay core.
- Heat flux obtained was a practical alternative in detecting seepage and water flux in an earth dam.

1. INTRODUCTION

Earthen dams are simple and cost-effective structures made of soil based on silicon that stand on their self-weight to prevent sliding and overturning. Earth dams are prone to seepage, and inadequate seepage may cause stability problems as it leads to a reduction in the shear strength of the dam, which could lead to its failure.

Radzicki & Bonelli (2010) applied an Impulse Response Function Thermal Analysis model to detect leakage in an earth hydraulic structure – an earth dam and dikes. Chen *et al.* (2009) developed a Finite element method-based model of fully coupled multiphase flow, thermal, and stress/deformation in a porous geological media. A three-dimensional software name 'THYME3D' was also developed. Johansson & Sjödahl (2004) inspected the downstream toe of about ten embankment dams to detect seepage with data obtained by measuring temperature. Yousefi *et al.* (2013) studied leakage in earth-fill dams using seepage flow and thermal simulation. Convective diffusion equation and mass balance were discretized using the finite element analysis in saturated and unsaturated zones. Results showed that temperature variation was more effective for inspection of leakage than the piezometric seepage levels. Alekseevich & Sergeevich (2017) developed a thermal-seepage regime numerical model based on the finite element method for dam foundations operated in permafrost conditions. Several factors that affect the temperature measurement were the thermal conductivity of soil, microclimate,

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (http://creativecommons.org/licenses/by/4.0/).

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

and the vegetation in the area. It was found that the amount of water rising can be estimated by computing the amount of heat energy transferred to the surface.

Abu-Hamdeh (2014) studied the effect of water content and bulk density on the specific heat and volumetric heat capacity in a laboratory on sand and loam. The calorimeter method was used to determine the specific heat of both types of soils. It was found that the specific heat increases with increasing moisture content of the soil. In contrast, the volumetric heat capacity increased with increasing the soil density and its moisture content. On similar moisture content and soil density, loamy soil had higher specific heat and volumetric heat capacity than the sandy soil. Amanifard *et al.* (2007) investigated the effect of an electrical double layer (EDL) near the solid/liquid interface on a three-dimensional heat transfer characteristic. Numerical investigation on the pressure drops of water flow through a rectangular microchannel was also carried out. It was observed that the liquid flow in rectangular microchannels was influenced significantly by the EDL, especially in the high electric potentials. Del Piero (2020) proposed a heat conduction theory in rigid heat conductors based on mechanical concepts and compared it with traditional thermodynamic theories. It was found that the proposed mechanical approach provided the same equation as the thermodynamic approach but in a simpler way. Badruddin *et al.* (2020) reviewed heat transfer in porous media in various geometrical shapes like a vertical plate, cylindrical shape, cavity, etc. Heat transfer in natural convection, mixed convection, thermal equilibrium, and thermal non-equilibrium was also explained for the review's purpose. Cuong *et al.* (2017) reviewed temperature measurement to estimate seepage in an embankment dam. The downstream toe was highlighted as the most important location to detect the seepage using fiber optic sensors. Short-term temperature analysis was good for estimating localized leakage, which helps minimize the cost of repair works. It was also reviewed that from the temperature measurements, seepage velocity can be evaluated.

Buntebarth (2020) determined the thermal conductivity and the thermal diffusivity of dry granular material by heat transfer between the source and the investigation media. An empirical relation was established between the thermal conductivity and the diffusivity of the media. Misra *et al.* (1995) presented a review on factors affecting soil thermal conductivity. They determined soil thermal conductivity by empirical and semi-empirical methods of five soil types: gravel, sand, silt, clay, and peat. A theoretical thermal conductivity model of granular soil material was developed in an almost dry phase. Abu-Hamdeh (2003) investigated the effect of water content and bulk density on the specific heat, volumetric heat capacity, and sandy and clay soil's thermal diffusivity. Laboratory experiments were performed using the calorimetric method. The results also observed that the clay soil generally had higher specific heat and volumetric heat capacity than sandy soil. Chuvilin & Bukhanov (2019) conducted an experimental study to determine the thermal conductivity of frozen soil at gas pressure below equilibrium using a KD-2 needle probe which caused a little impact on the soil samples in the study area.

Kurz *et al.* (2017) presented experimental laboratory results to estimate the thermal conductivity of frozen and unfrozen soil samples of clay, silt, and peat subjected to seasonal freezing and thawing at the study area. Thermal conductivity values obtained from empirical methods were compared with the value obtained from a thermal probe in the laboratory. Kodešová *et al.* (2013) measured thermal conductivity and heat capacity using a KD2 PRO device with 13 TR-1 and SH-1 sensors. The results measurement observed that the highest thermal conductivities were measured in soils on quartz and substrates.

2. SOIL PROCUREMENT AND ITS TESTING

The soil was procured from a depth of 2 m below the ground level of the ground of Delhi Technological University, Delhi, India (28.7501° N, 77.1177° E). The soil's geotechnical properties, used to construct the dam in hydraulic flume, were tested in the laboratory. The test results for geotechnical properties were used to develop the dam model, such as natural moisture content, sieve analysis, liquid limit, plastic limit, permeability, standard compaction test, consolidation test, and direct shear test. X-ray diffraction (XRD) test was carried out using Bragg-Brentano with a Bruker D8 advance machine to find out the crystalline materials in the soil used in the construction of the shell and core of the earth dam models in the hydraulics laboratory.

3. THERMAL PROPERTIES

3.1. Thermal conductivity

In the past, soil thermal properties were used to predict water, heat, and solute transport in soils (Abu-Hamdeh 2003). Among other essential thermal properties, soil thermal conductivity has a crucial role in temperature modeling in a porous medium. It depends on moisture content, mineral composition, temperature, and texture.

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

However, for low-density clay soils, the critical moisture content coincides with the plastic limit. The thermal conductivity of coarse-textured, angular grained soils is higher than that of fine-textured soils (Misra *et al.* 1995). Various authors have reported the thermal conductivity of soil at different temperatures and mineral compositions. For silty sand having quartz content of 64%, water content of 15%, and dry density of 1.77 g/cm^3, reported thermal conductivity of 1.04 W/mK in the unfrozen state and 1.61 W/mK in the frozen state at −6 °C. (Chuvilin & Bukhanov 2019). The thermal conductivity of sand collected from two places had a value of 1.9, 1.25 at 20% water content. The thermal conductivity of saturated sand was 3.12 W/mK (Buntebarth 2020). The average water thermal conductivity value W/mK was 0.56 at 17 °C (Chuvilin & Bukhanov 2019). The thermal conductivity of water was reported to be 0.594 W/mK (Kosky *et al.* 2012).

3.2. Specific heat and volumetric heat capacity

'Specific heat is the amount of heat required to change the temperature of a mass unit of a substance by one degree. In other words, the specific heat capacity of a substance is the heat capacity of a sample of the substance divided by the sample's mass. It is used to calculate the energy change associated with a temperature change.' Many factors influence soil heat capacity, among which water content and soil density are the key factors that influence it. In sandy soil, specific heat is associated with water content for known bulk density. Water in the liquid state has one of the highest specific heats, comparing with common substances. The specific heat of soil is determined using a calorimeter by measuring its heat capacity and dividing its mass (Kosky *et al.* 2012). The sum of volumetric heat capacities of each soil component multiplied by their fraction gives soil volumetric heat capacity. 'It stands for stored internal energy ability of a given soil volume while undergoing a given temperature change.' The analysis of coupled heat-water transfer in soil plays a vital role in the temperature and heat transfer within an earth dam. Various literature has mentioned specific heat and volumetric heat capacity, among which few are presented. Dry sand had specific heat of 932–958 J/kgK measured using a calorimeter (Ižvolt & Dobeš 2014). While specific heat of the sand, quartz, is 830 J/Kg.°C, dry soil is 800 J/Kg.°C and wet soil is 1,480 J/Kg °C, respectively. (Engineering toolbox 2003, Specific heat of some common substances). Kodešová *et al.* (2013) reported a soil's specific heat of 0.73 kJ/KgK, and a volumetric heat capacity of 1.9 MJ/m^3/K, at the dry density of 1.34 g/cm^3. Specific heat of the water was 4.18 kJ/kgK and volumetric heat capacity was 4.18 MJ/m^3/K. For clay of dry density 1.46 g/cm^3, specific heat was 0.86 kJ/kgK and volumetric heat capacity was 2.30 MJ/m^3/K. While specific heat of water was 4,190 kJ/kgK. And at 30 °C, the specific heat of the water was recorded as 4.1175 kJ/kgK or 74.181 J/(mol k) (Engineering toolbox 2004, water - specific heat).

4. HEAT TRANSFER IN SOIL

'Heat transfer refers to thermal energy that passes between objects by either thermal conductivity, thermal convection, in which a fluid moves between regions of different temperature, or thermal radiation, in which energy is transmitted by electromagnetic radiation.' Temperature difference induces a heat flux. This induced heat flux flows from the hotter medium to the colder; it requires a difference in the temperature and a medium through which heat flows. Heat flux also depends on the thermal transfer coefficient of soil used in constructing the earth dam and its foundation. In an earth dam, heat can flow through soil-to-soil particles (conduction) and fluids' motion in the pores (convection). The air temperature surrounding the earth dam and water in the reservoir upstream of the earth dam serves as the dam's principal thermal loading. Other sources that may affect the earth dam's temperature variation are geothermal, a frozen process in colder regions, humidity changes around an earth dam, and wind influence. Tokoro *et al.* (2016) proposed a thermal conductivity model and empirical equations for soil. The thermal conductivity of soil and its moisture content had a non-linear relationship. Ižvolt & Dobeš (2014) monitored the impact on the measured value of three different building materials' specific heat and thermal conductivity due to the calorimetry test procedure's modification. Alim *et al.* (2017) used Seep/w to calculate seepage discharge for a homogeneous earth dam resting on an impervious base without any filter or clay core. They proposed an equation for seepage through a homogeneous earth dam resting on an impervious base.

4.1. Heat conduction

'It is defined as the amount of heat passing in unit time through a unit cross-sectional area of the soil under a unit temperature gradient.' Viscosity and water density are temperature-dependent phenomena in the earth dam,

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

which affect the heat conduction in the soil particles present in an earth dam's body. Heat conduction can occur due to soil's material property present in the earth dam body and foundation. Heat conduction is affected by the saturated/unsaturated region in the earth dam, as below the phreatic line, natural or free convection also occurs. Above the phreatic line, conduction predominates; that is, in the upper zone of the dam, there is pure heat conduction and no heat advection as negligible water flow is present above the phreatic line; hence, slow heat transport from the dam surface inward to the dam occurs. While in the central zone, both convection in x-direction and conduction in y-direction takes place.

4.2. Convection in soil

Convective heat transfer occurs in an earth dam whenever there is a difference in temperature in the earth dam body and the water upstream. Natural convection is produced by motion caused due to density difference in a fluid due to temperature difference, called convection currents. It continues as long as temperature difference exists in an earth dam and upstream water. Temperature acts as a natural tracer to find the seepage flow and water flux to localize leakage in embankment dams. In an earth dam, heat flux was used to detect leakage regions and determine seepage velocity using temperature observation. When seepage water goes through an earth dam's pores, it exchanges heat with the adjacent soil medium (Cuong *et al.* 2017). Heat flux varies in an earth dam due to convective heat transfer, due to the energy transported by seepage flow. Increased fluid speed enhances high heat transfer in convection, and with high heat transfer, a high-pressure drop can be observed.

5. METHODOLOGY

The physical model was scaled down so that it was possible to perform experiments in the laboratory using a hydraulic flume. The length scale factor for the experimental model was 1/100. This scale factor was applied to the length, width, and height of the dam, and the discharge obtained. The effect of the longitudinal slope was studied by keeping these dimensions: upstream slope 38°, downstream slope 45°, bottom width 52 m; the height of the dam was kept at 20 m, while the longitudinal slope was varied at 0.00%, 6.07%, 2.28%, 4.00% slope for the models 1,6,7,8 respectively. The effect of the central impervious core (clay) on water and heat flux was studied in models 9 and 10 by comparing with the homogeneous model (model 1). The dimensions of models 9 and 10 were similar to model 1 but with an impervious core of bottom width 18 m each and top width of 2.4 m and 1.9 m, respectively. Table 1 shows the different dimensions of the homogeneous earth dam models with their dimensions and upstream head. Table 2 shows an earth dam's models with different horizontal filter lengths and a filter thickness of 3 m.

5.1. Experimental setup

An earthen dam of known dimensions was scaled down for experimental work; 16 earth dam models were constructed in a hydraulic flume of dimensions 300 cm length, 30.5 cm width, and 40 cm height. A Proctor penetrometer was used to control the compaction in layers of the earth dam during construction. The effect of scaling was considered as it affects the outcome of the result. A homogenous as well as an earth dam with central clay core was constructed. The temperature variation inside the dam was measured to calculate the heat flux across the earth dam in each model due to the convective water flow. Hydraulic flume was titled to change

Table 1 | Earth dam models' dimensions and upstream head of homogeneous section

Model number	Upstream water head (m)	Upstream slope (degrees)	Downstream slope (degrees)	Bottom width of the dam (m)	Height of dam (m)
1	13.5	38	45	52	20
2	13.5	38	59	52	20
3	13.5	50	34	52	20
4	13.5	50	34	56	20
5	13.5	50	34	56	15
11	13	38	45	52	20
12	13.5	45	45	52	20
13	13.5	50	45	52	20

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

Table 2 | Earth dam models with different filter lengths

Model number	Upstream slope (degree)	Downstream slope (degree)	Bottom width of dam (m)	Height of dam (m)	Length of the filter from the downstream end (m)	% length ratio of filter to the total length (m)
1	38	45	52	20	0.0	00.00
14	38	45	52	20	15.5	29.80
15	38	45	52	20	23.0	44.23
16	38	45	52	20	13.0	25.00

the longitudinal slope to study its effect on water flux and heat flux. Fluid viscosity was adjusted by adding sugar and then measuring the rotational rheometer, a viscosity measurement instrument, in the physics laboratory to study its effect on water flux and heat flux.

Proper compaction in an earth dam leads to increased stiffness and strength, minimizing the settlements and preventing the dam's sliding failure. A Proctor penetrometer was used to achieve proper compaction in the construction of the 16 models of the earth dam. The soil was placed in five layers of 4 cm each at optimum moisture content with 90–95% variance using a Proctor penetrometer while keeping each layer's penetration rate at 1.3 cm/s. An average of three penetration resistance were compared with the calibration chart of moisture-penetration resistance and moisture-density curve relationship. A longitudinal filter was provided, and its length was varied for studying its effect on water flux and heat flux. Models 14–16 were constructed with a downstream filter with varying lengths, as mentioned in Table 2. In the present study, the water flux was obtained from the seepage discharge calculation; while heat flux in models was calculated using temperature readings obtained by a digital thermometer in a hydraulic flume.

5.2. Numerical modeling

Geostudio 2020 software is one of the robust programs based on the finite element technique and can simulate and analyze water flux using Seep/w and heat flux using Temp/w. A finite element mesh of size 0.1 m was chosen after several trial computations by changing the mesh size. The mesh size was determined so that it doesn't affect the computation time to a large extent and yields a good result.

6. RESULTS AND DISCUSSION

- The basic properties of soil were tested, and their results are shown in Table 3.

Figure 1 showing results of intensity vs. angle (2Θ) for soil used in the construction of homogeneous earth dam models as well as in the shell of the earth dam in non-homogeneous earth dam models, while Figure 2 shows XRD results of intensity vs. angle (2Θ) for soil used in the core of non-homogenous earth dam. Silicon oxide

Table 3 | Basic properties of soil

S.No.	Name of the test		Result
1	Natural moisture content		1.49%
2	Specific gravity		2.57
3	Grain size distribution	Soil classification	Silty sand
		D_{60}	1.61
		D_{30}	0.36
		D_{10}	0.09
		C_u	17.19
		C_c	0.87
4	Atterberg's limit	Liquid limit	26.24%
		Plastic limit	12.69%
5	Compaction:	Dry unit weight of the soil	18.05 kN/m^3
		Optimum moisture content	13.10%
6	Permeability of soil (shell)		6.78×10^{-6} m/s

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

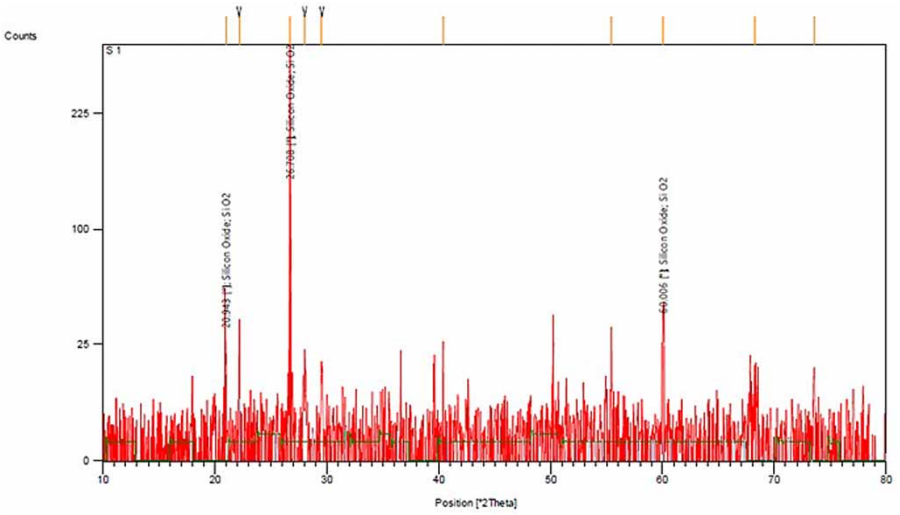


Figure 1 | XRD of soil used in the shell of earth dam models.

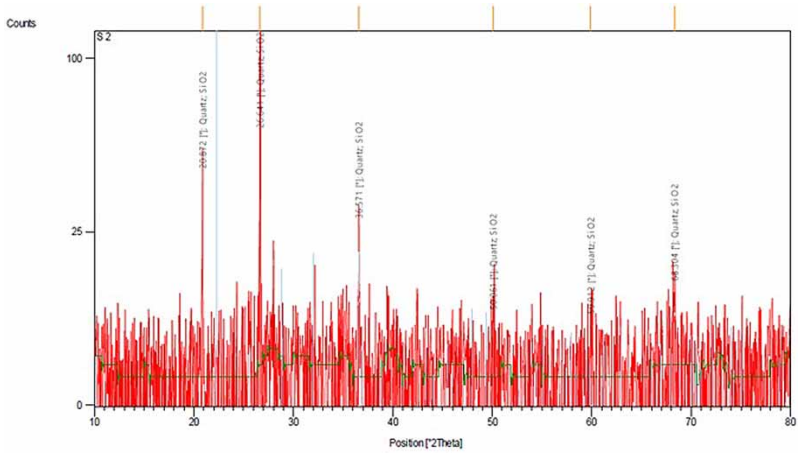


Figure 2 | XRD of soil used in the core of the earth dam models.

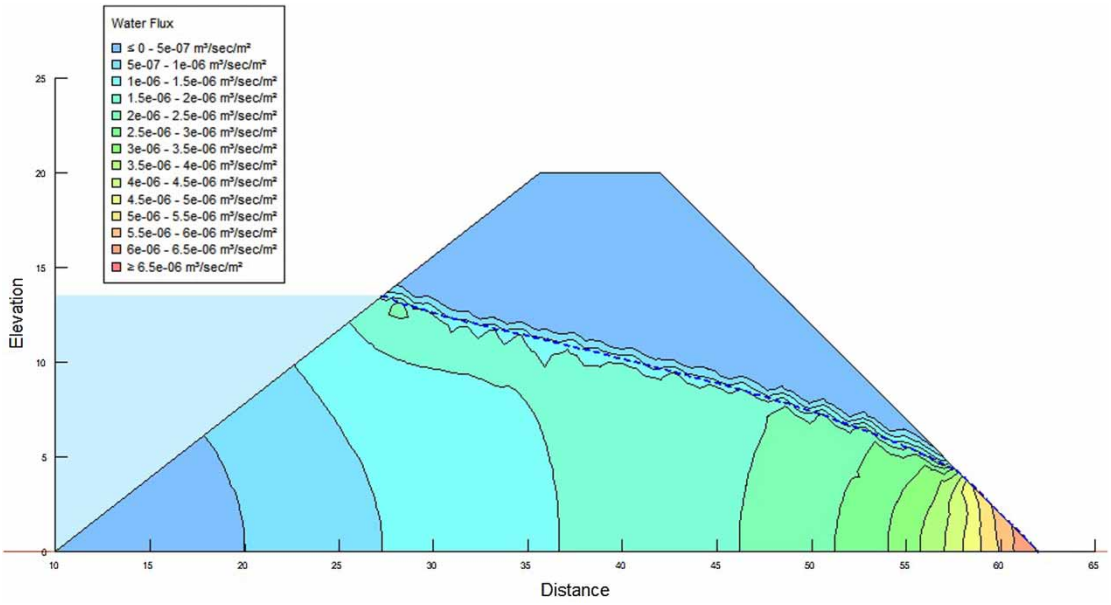


Figure 3 | Water flux in model 1.

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

(SiO_2) was the primary chemical compound in the soil (shell) as the angle (2Θ) was at 26.715 with an intensity of 100% and having the empirical formula of O_2Si. The h, k, and l were 0,1, and 1 respectively, and spacing between planes (d) equals 3.33420 A°. In contrast, quartz was the primary chemical compound found in the soil (core), having the chemical formula SiO_2 and empirical formula O_2Si. The lattice's crystal system was hexagonal with an angle (2Θ) of 26.64 at an intensity of 100%. h, k, l, of the lattice was 0,1,1 respectively, and spacing between planes (d) equals 3.34353 A°.

6.1. Contours of water flux and heat flux

The contours of water flux in $m^3/sec/m^2$ are shown in Figures 3–18 for models 1 to 16, while heat flux is shown in Figures 19–34. The variation of water flux and heat flux for the location is shown. The earth dam model's phreatic

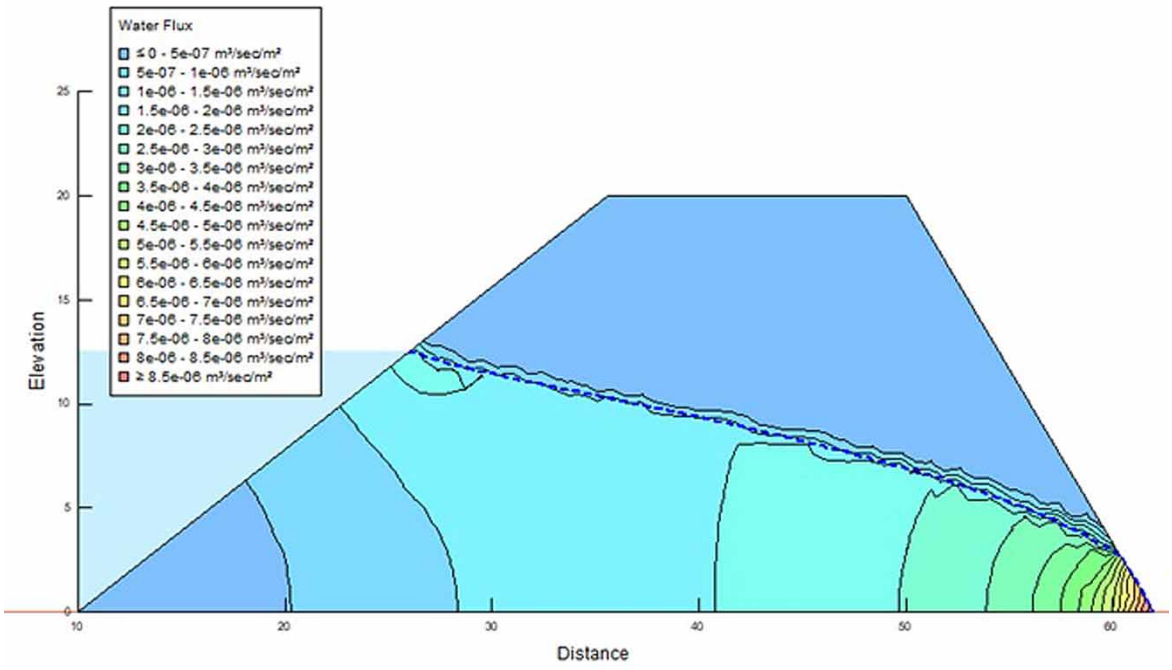


Figure 4 | Water flux in model 2.

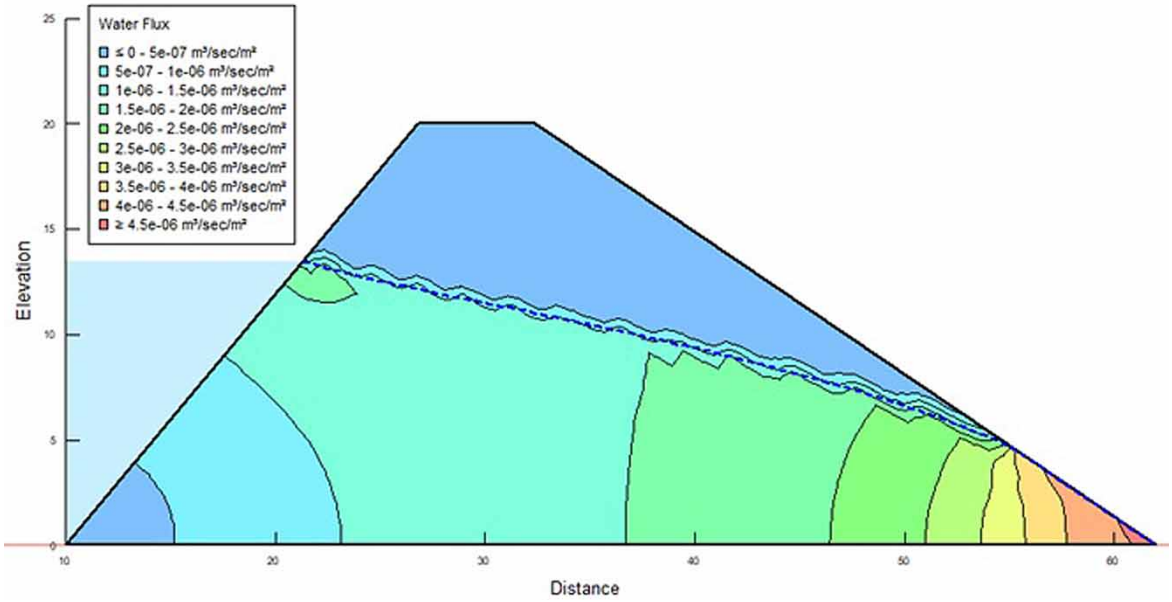


Figure 5 | Water flux in model 3.

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

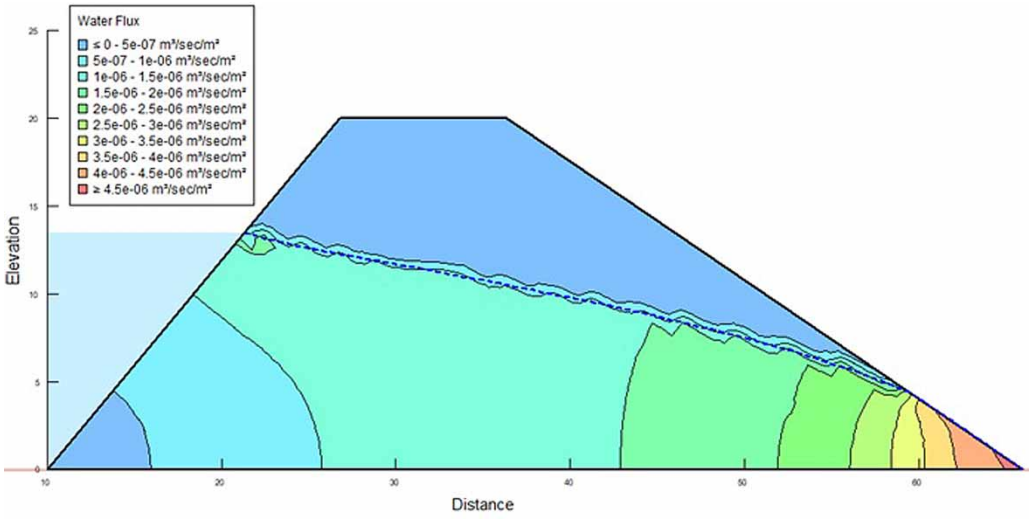


Figure 6 | Water flux in model 4.

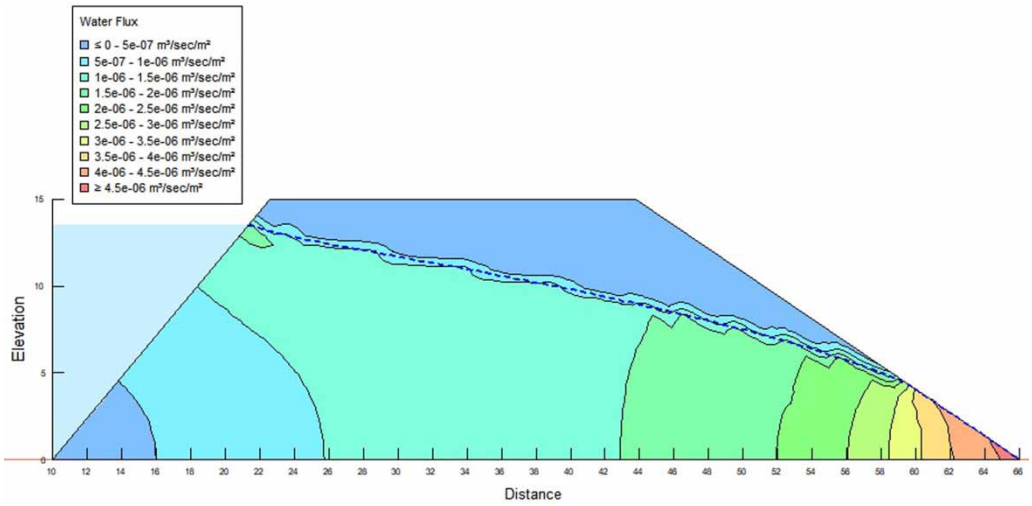


Figure 7 | Water flux in model 5.

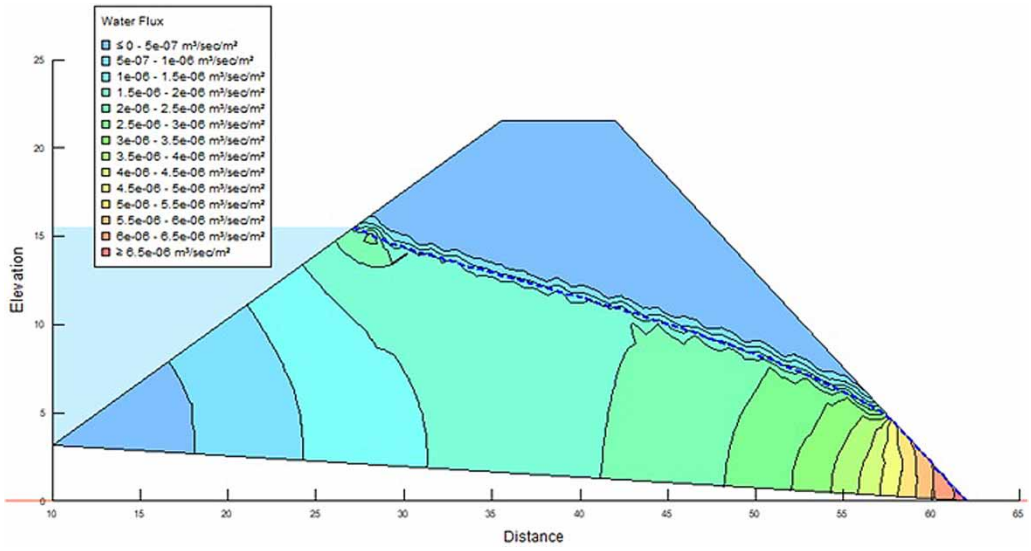


Figure 8 | Water flux in model 6.

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

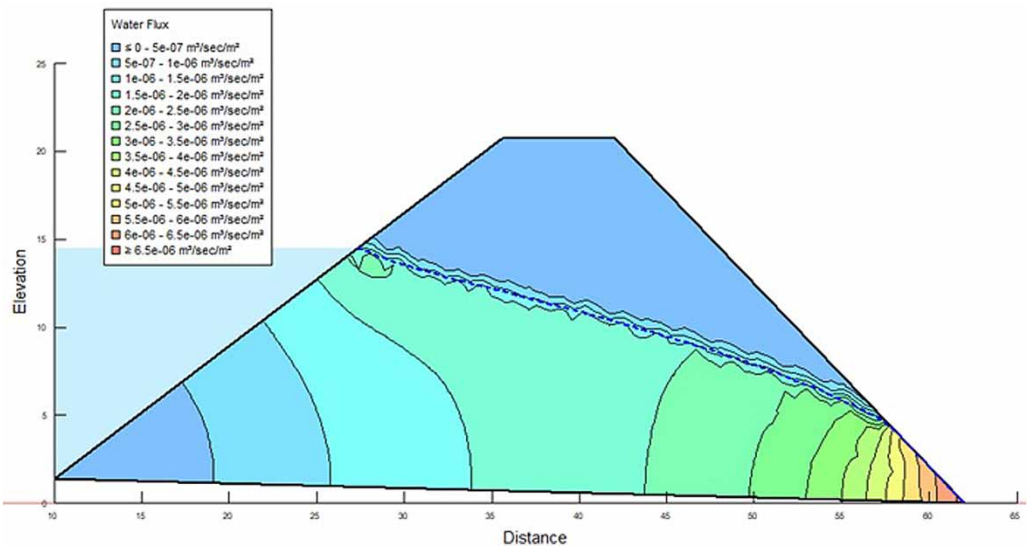


Figure 9 | Water flux in model 7.

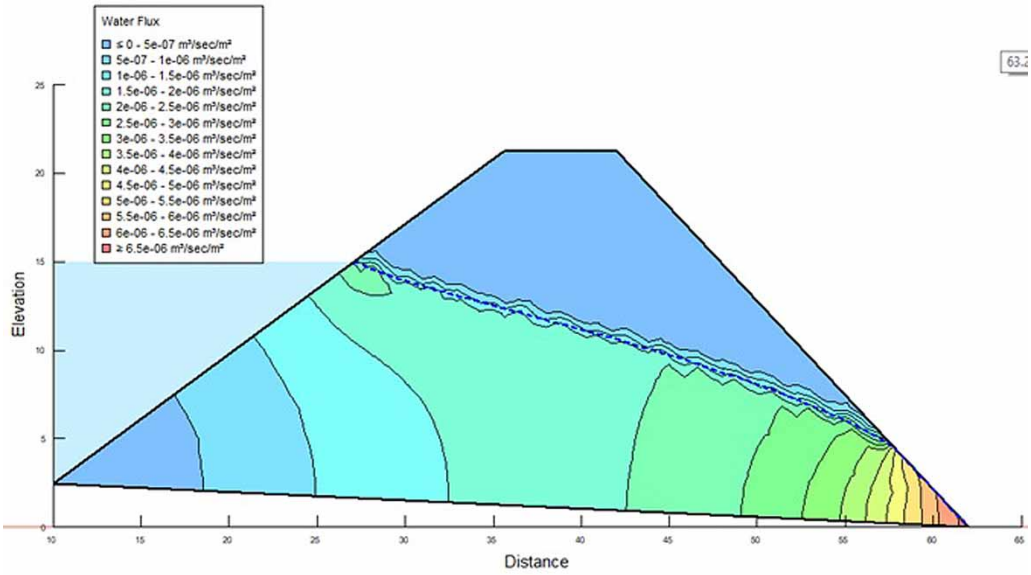


Figure 10 | Water flux in model 8.

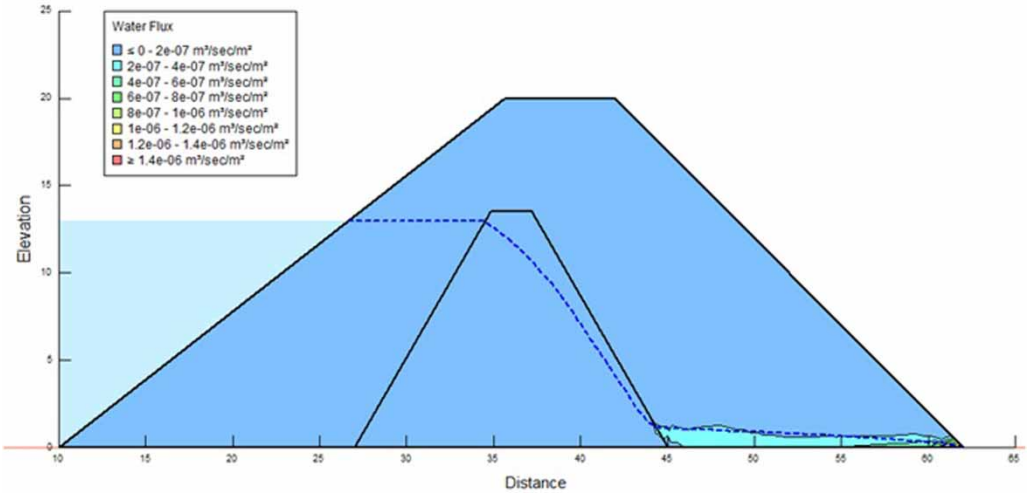


Figure 11 | Water flux in model 9.

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

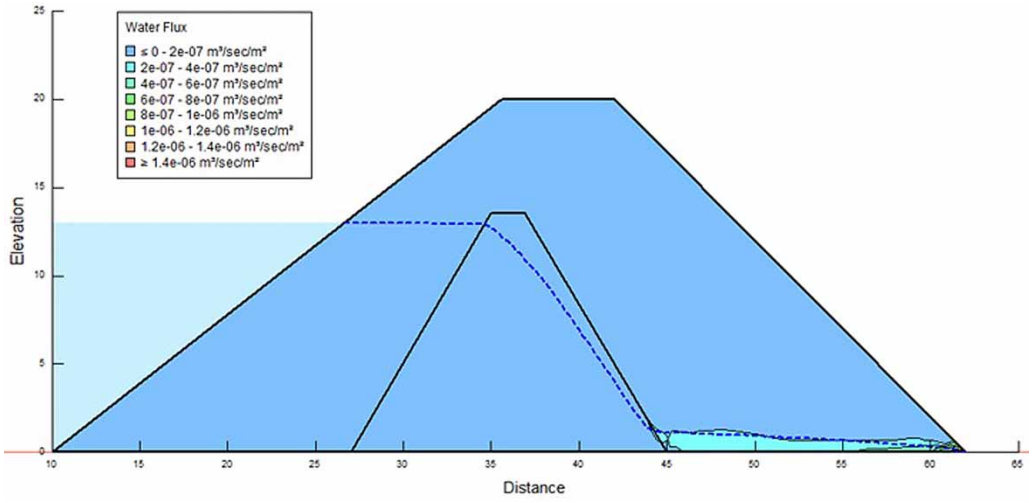


Figure 12 | Water flux in model 10.

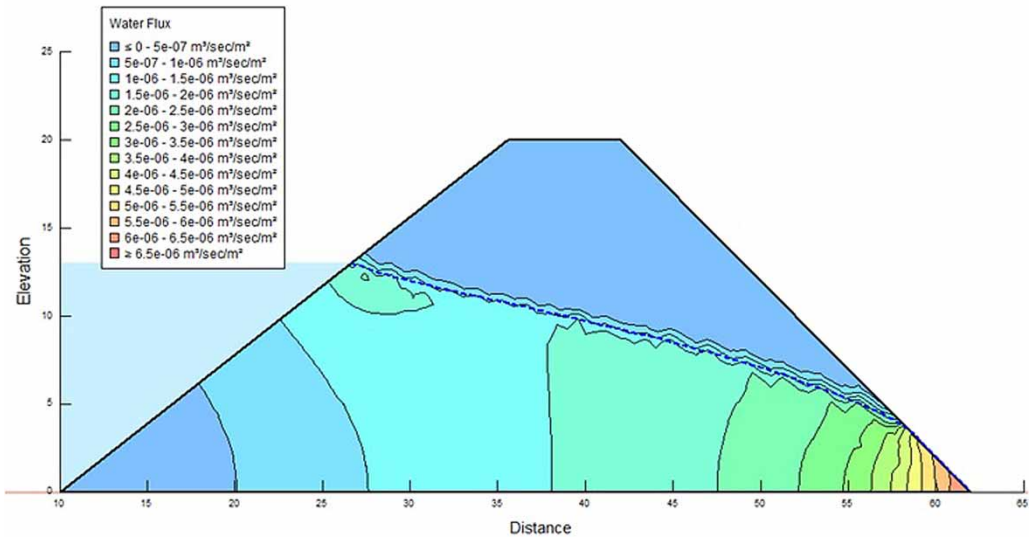


Figure 13 | Water flux in model 11.

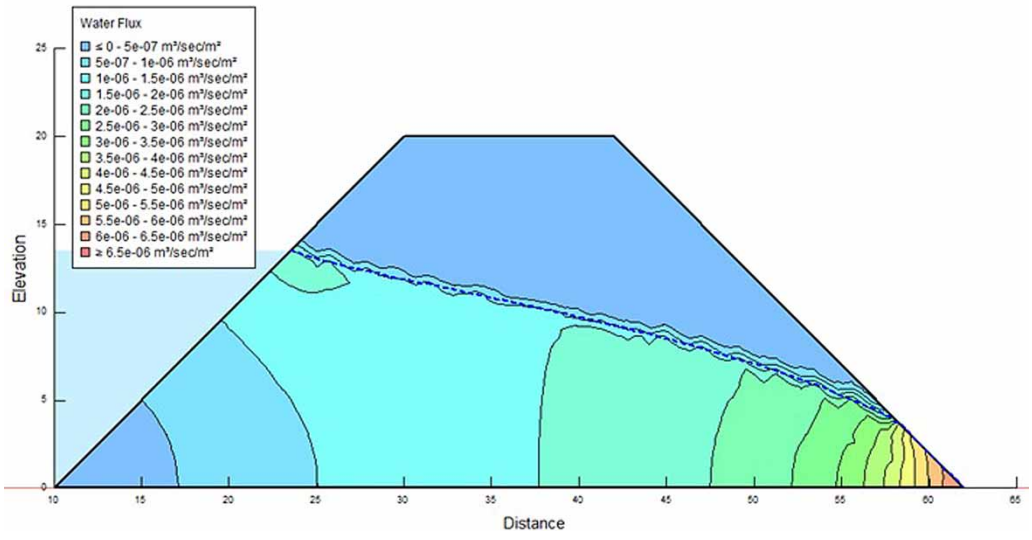


Figure 14 | Water flux in model 12.

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

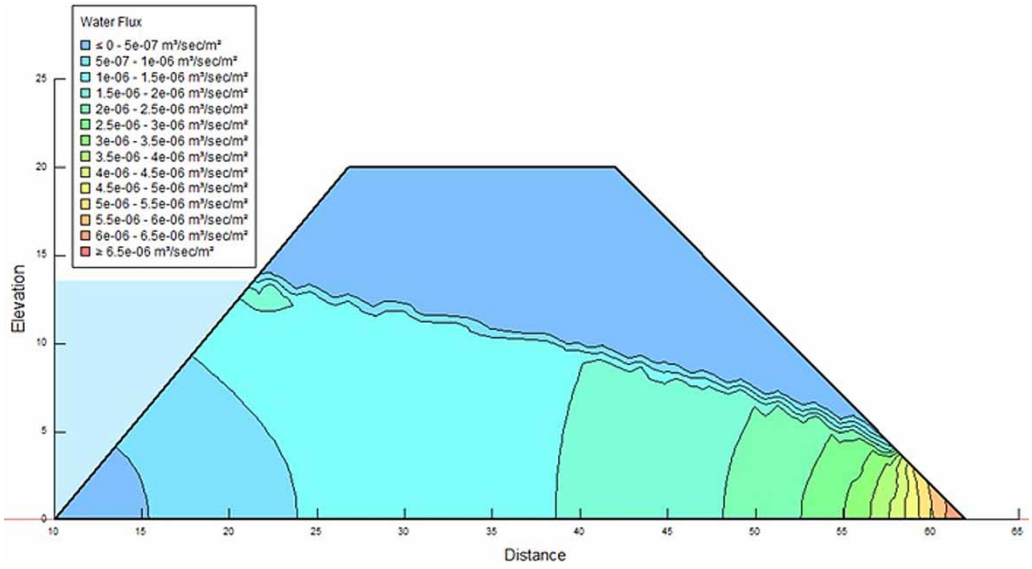


Figure 15 | Water flux in model 13.

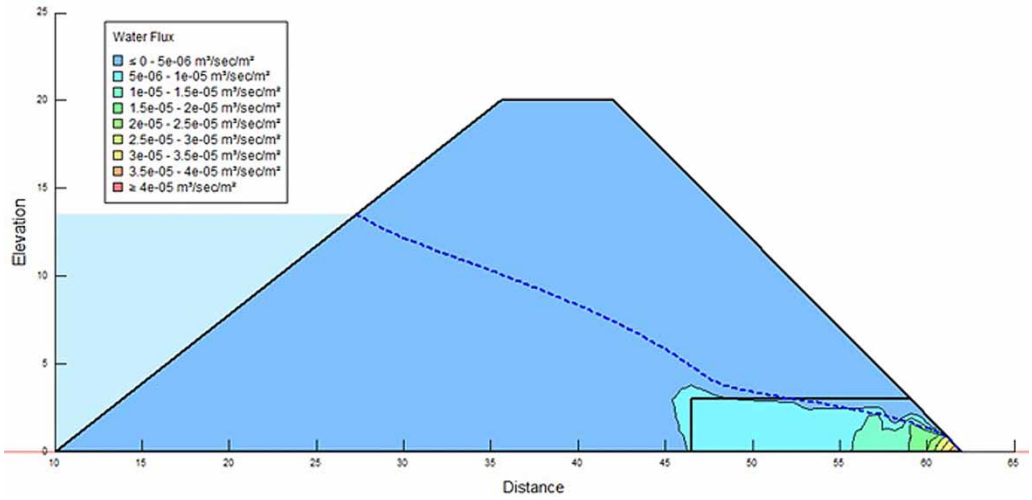


Figure 16 | Water flux in model 14.

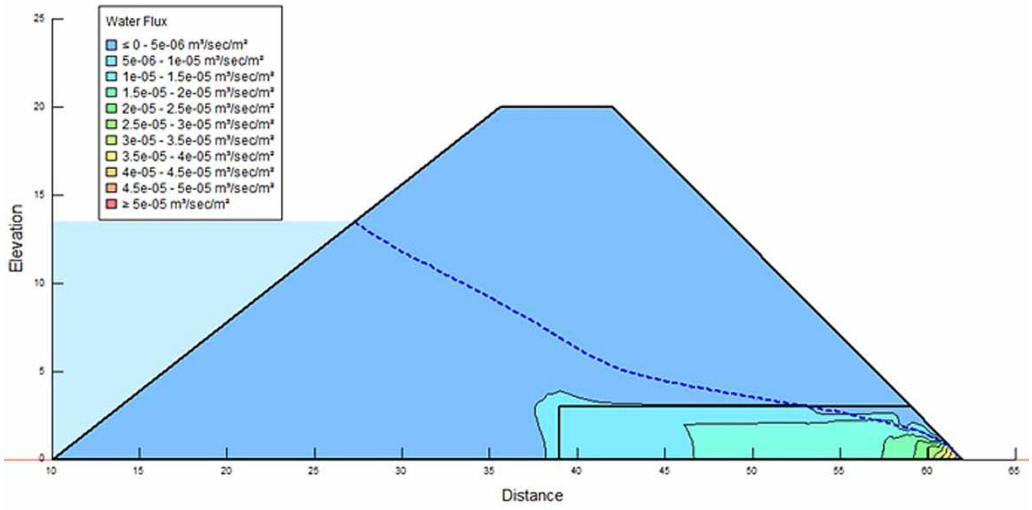


Figure 17 | Water flux in model 15.

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

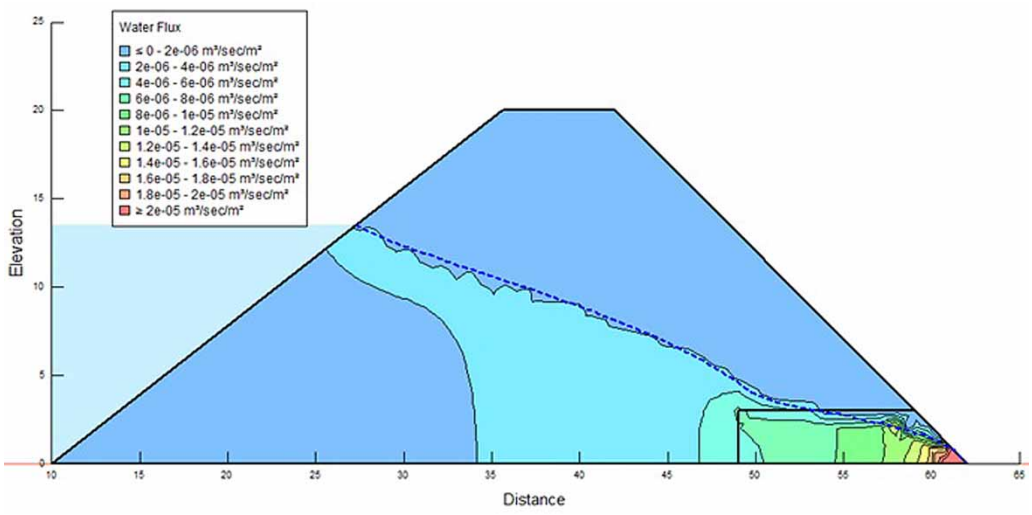


Figure 18 | Water flux in model 16.

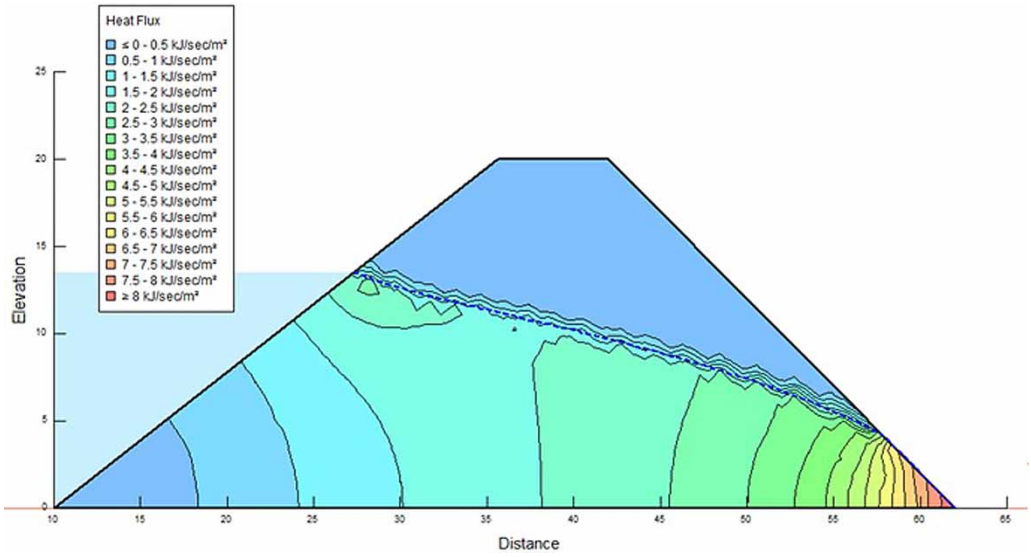


Figure 19 | Heat flux in model 1.

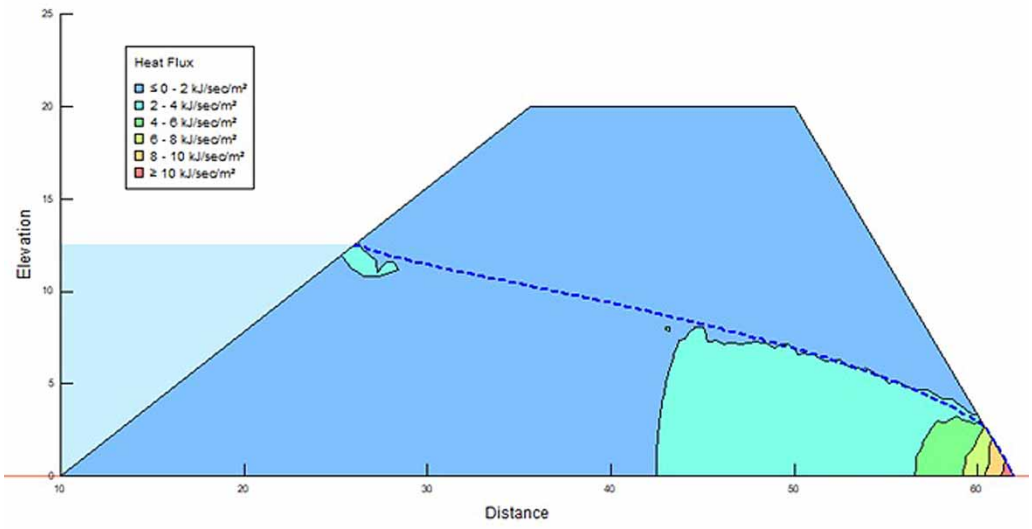


Figure 20 | Heat flux in model 2.

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

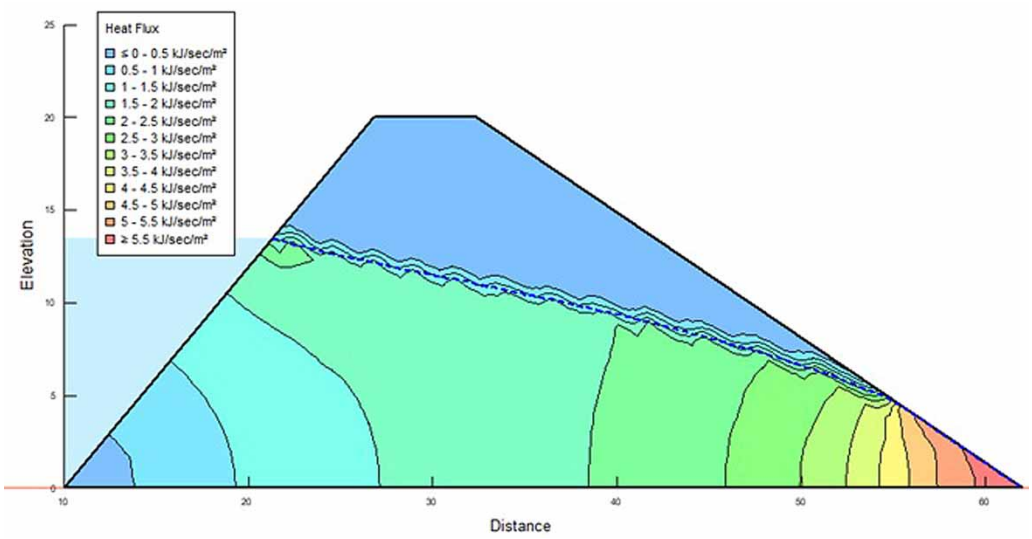


Figure 21 | Heat flux in model 3.

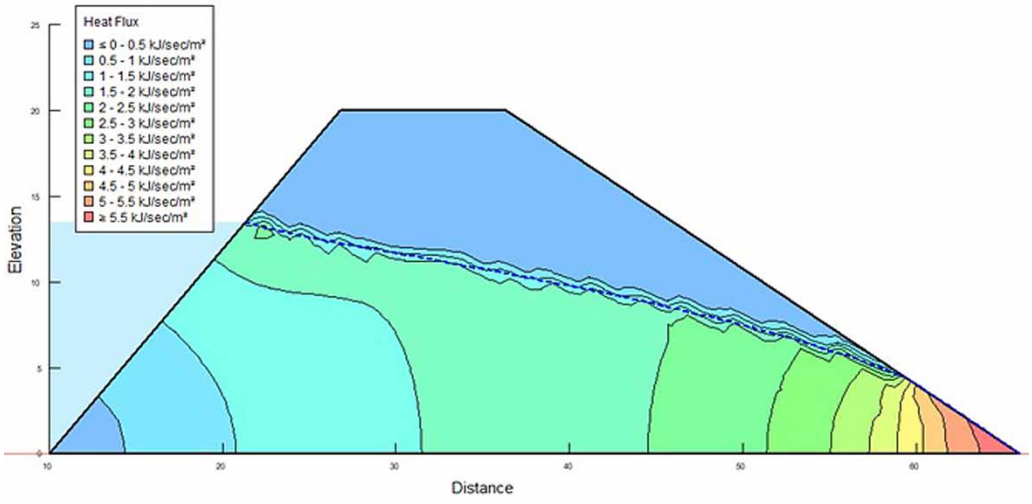


Figure 22 | Heat flux in model 4.

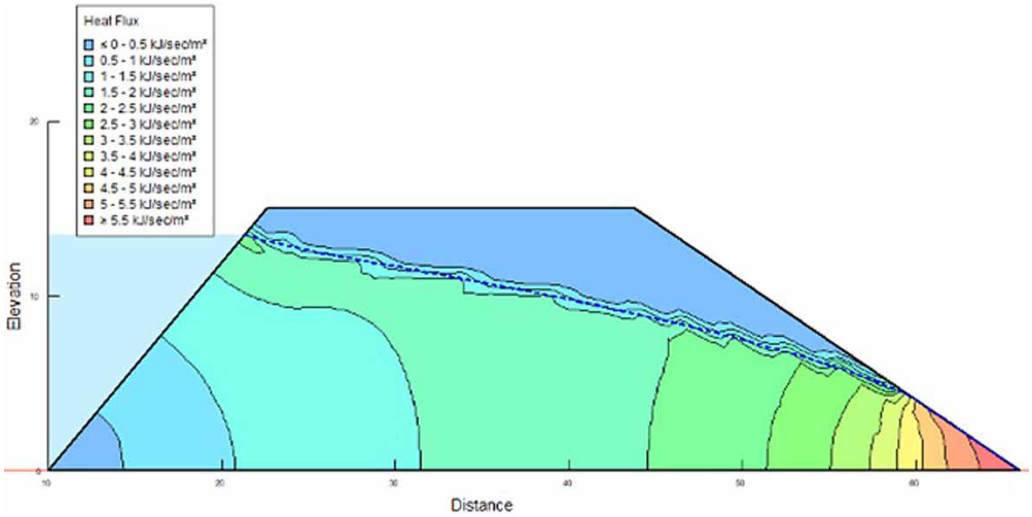


Figure 23 | Heat flux in model 5.

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

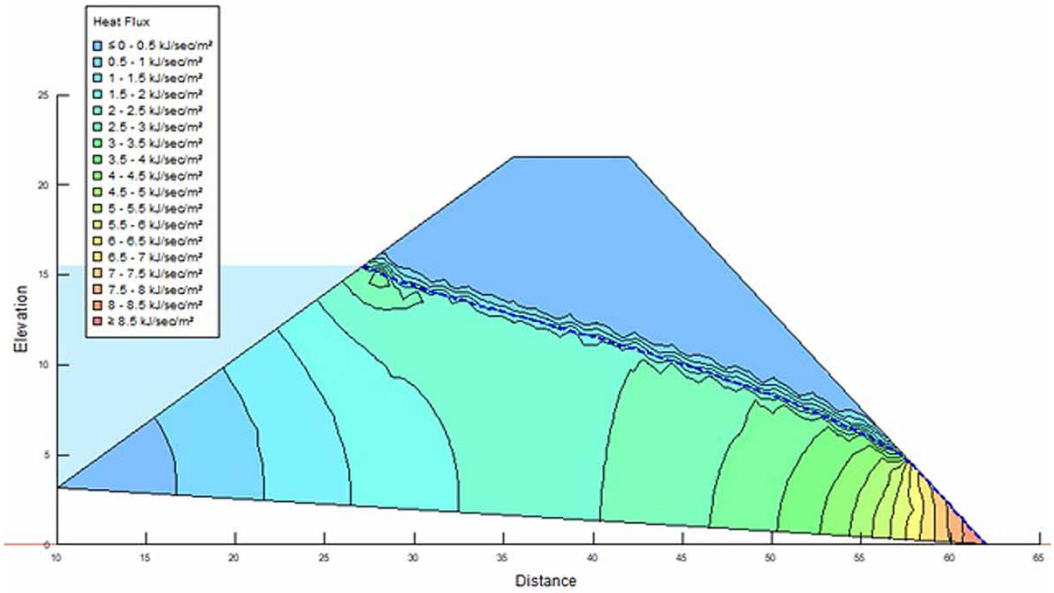


Figure 24 | Heat flux in model 6.

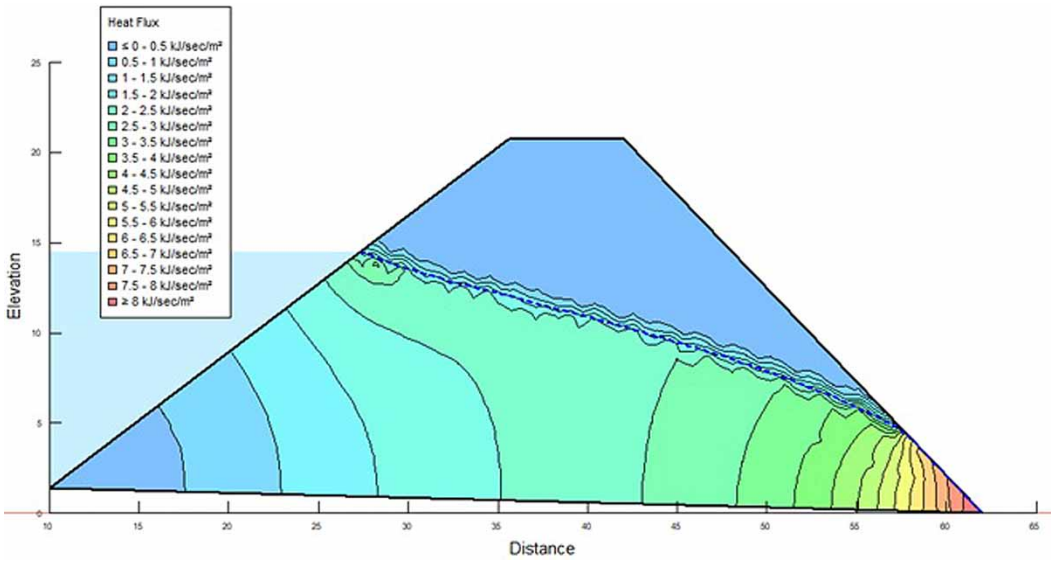


Figure 25 | Heat flux in model 7.

line marked with a dotted line in its respective figure is also shown, which keeps the flow of heat and water below it.

- Figure showing the graph of heat flux is Figure 35, with its x-axis showing the number of earth dam models, and the y-axis denoting heat flux in kJ/sec/m^2. Figure 36 represents a water flux graph with its x-axis as the number of earth dam models, and the y-axis denotes water flux in m^3/sec/m^2. While Figure 37 shows the flux on the y-axis to compare heat flux and water flux with the respective model shown on the x-axis.
- On increasing the downstream slope of an earth dam from 45 degrees (in model 1) to 59 degrees (in model 2), the water flux and heat flux increased by 34.81% and 34.37%. While increasing the downstream slope from 34 degrees (in model 3) to 45 degrees (in model 13), the water and heat flux increased by 42.36%.
- On increasing the upstream slope from 38 degrees (in model 1) to 45 degrees (in model 12); that is, 18.42%, the water flux and heat flux decrease by 0.34% and 0.99%, respectively. While increasing the slope from 38 degrees (in model 1) to 50 degrees (in model 13); that is, 31.57%, water flux and heat flux both decreased by 0.51%.

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

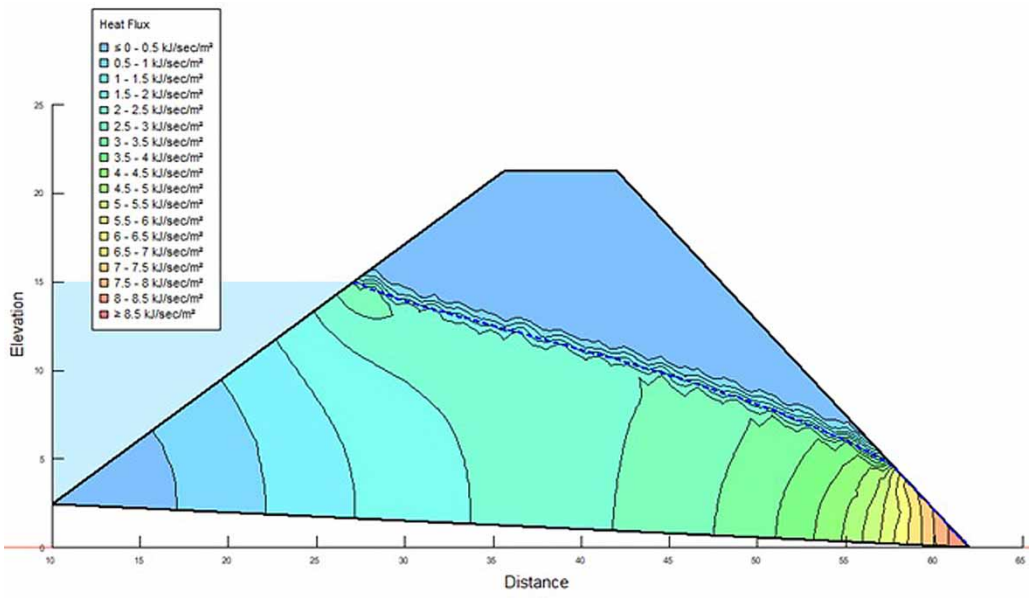


Figure 26 | Heat flux in model 8.

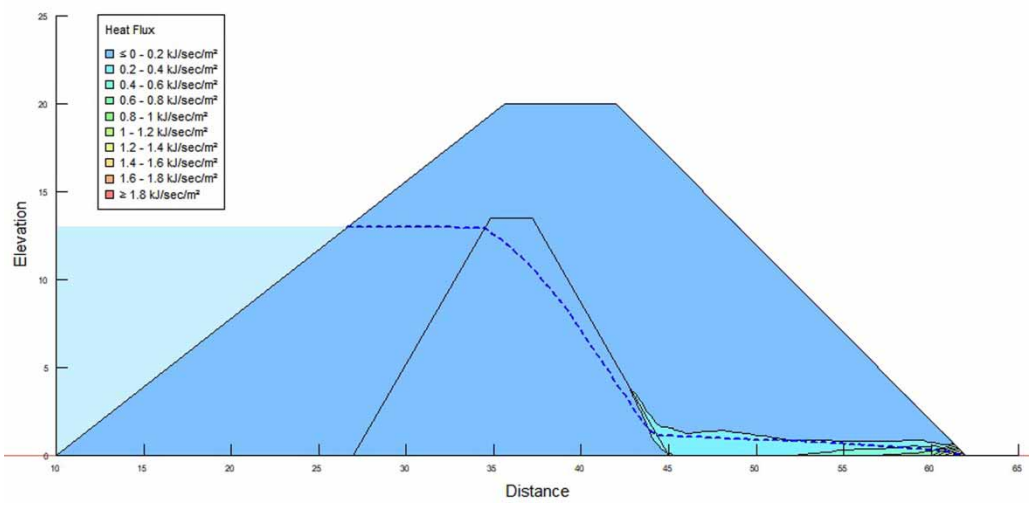


Figure 27 | Heat flux in model 9.

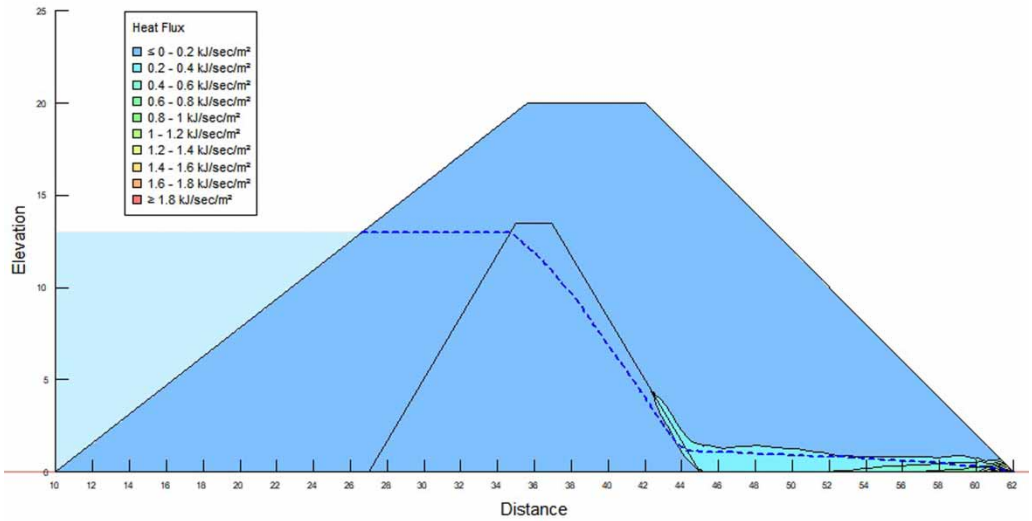


Figure 28 | Heat flux in model 10.

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

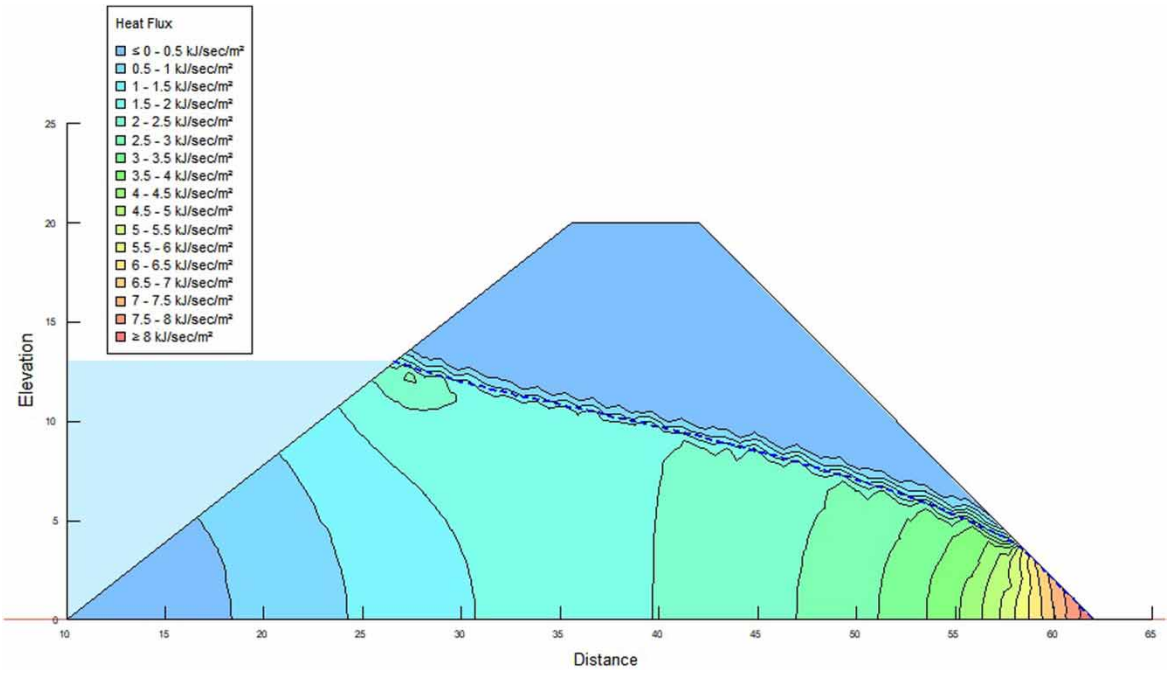


Figure 29 | Heat flux in model 11.

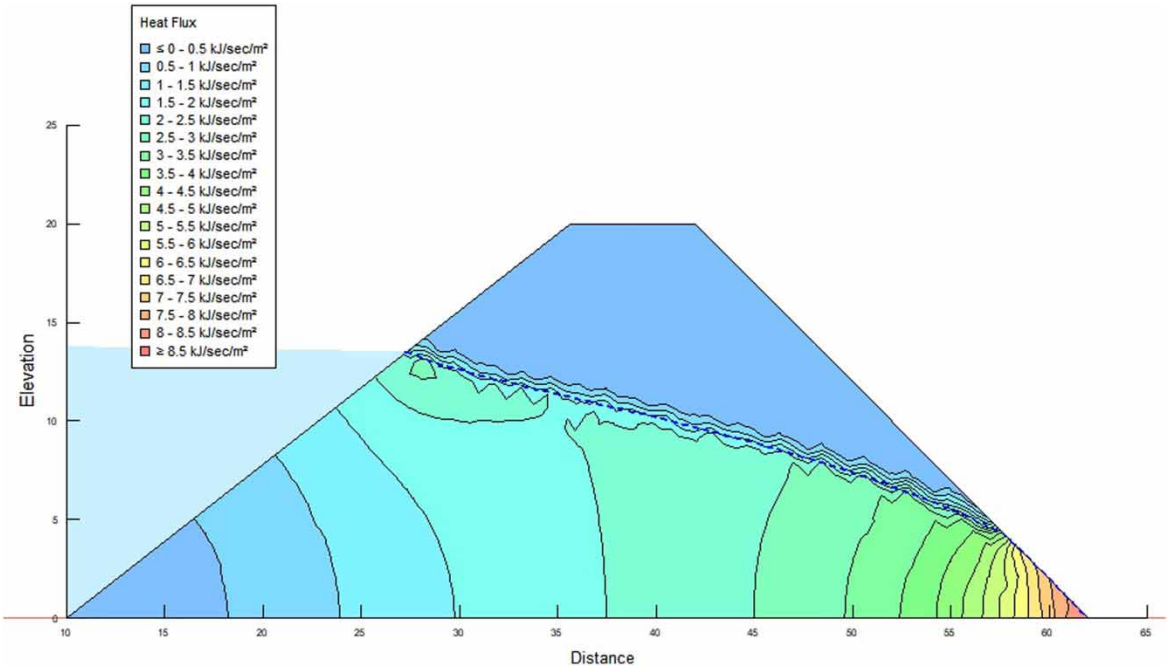


Figure 30 | Heat flux in model 12.

- On increasing the dam's length from 52 m (in model 3) to 56 m (in model 4); that is, by 7.69%, the water flux and heat flux both had an insignificant reduction of 0.02%.
- On increasing the earth dam's height for the same upstream head by 33.33%; that is, 15 m (in model 5) to 20 m (in model 4), the water flux and heat flux remain unchanged.
- On increasing upstream head of water from 13 m (in model 11) to 13.5 m (in model 1); that is, by 3.85%, the water flux and heat flux increased by 0.34%.

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

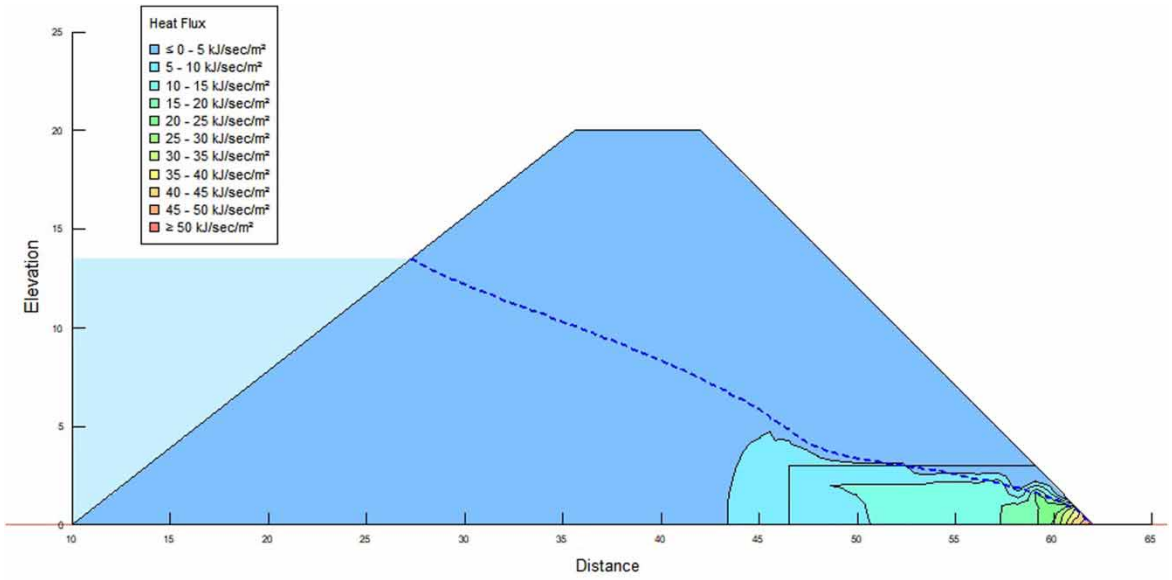


Figure 31 | Heat flux in model 13.

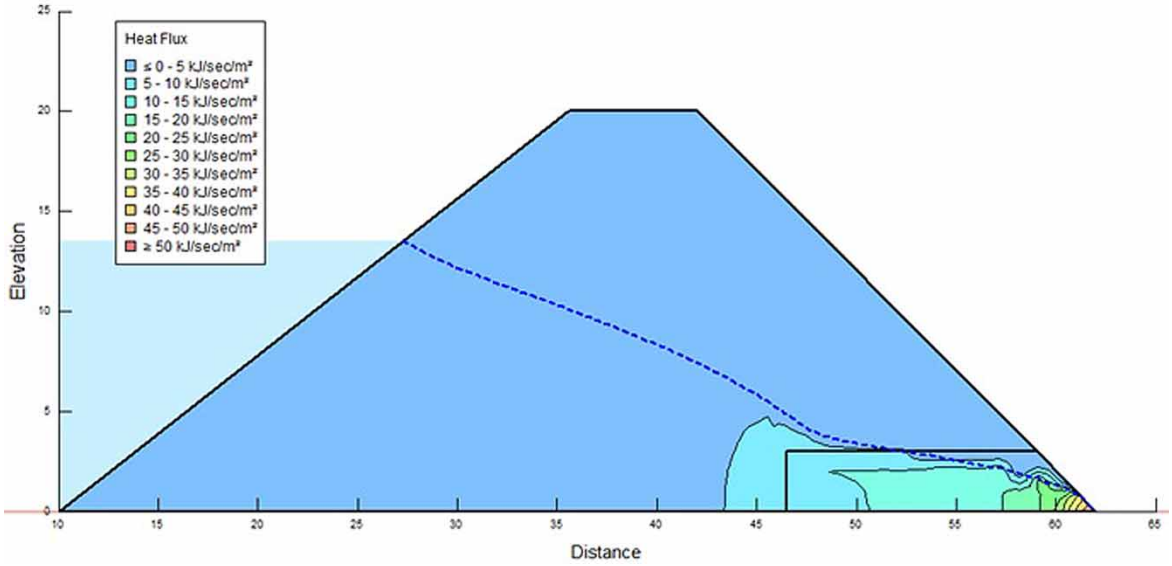


Figure 32 | Heat flux in model 14.

- On increasing the longitudinal slope by 2.28% (in model 1 to model 7), water flux and heat flux both increased by 1.33%, and increasing longitudinal gradient by 4% (in model 1 and 8), water flux and heat flux increased by 2.01% and 2.06% respectively. While increasing slope by 6.07% (from model 1 to model 6), both water flux and heat flux increased by 2.55%.
- With a decrease of the top width of clay core by 20.83% from model 9 to model 10, both the water and heat flux increase by 3.44%.
- With the downstream filter of length 13 m in model 16, 15.5 m in model 14, and 23 m in model 15, water flux increased by 414.23%, 533.38%, and 666.51%, respectively, and heat flux increased by 414.23%, 533.38%, and 666.75% respectively.

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

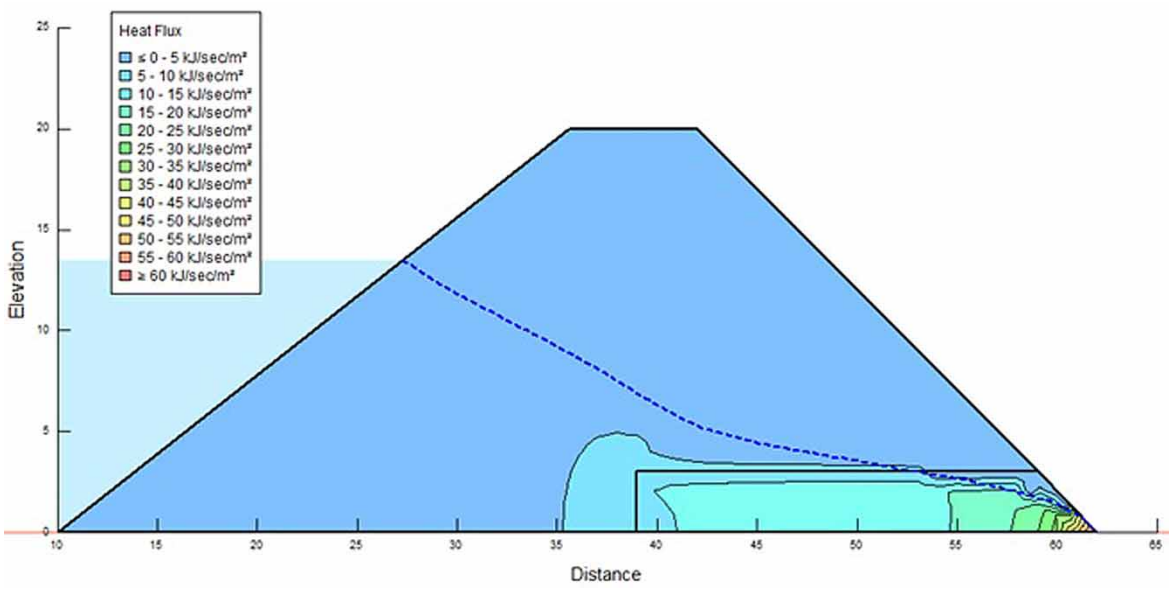


Figure 33 | Heat flux in model 15.

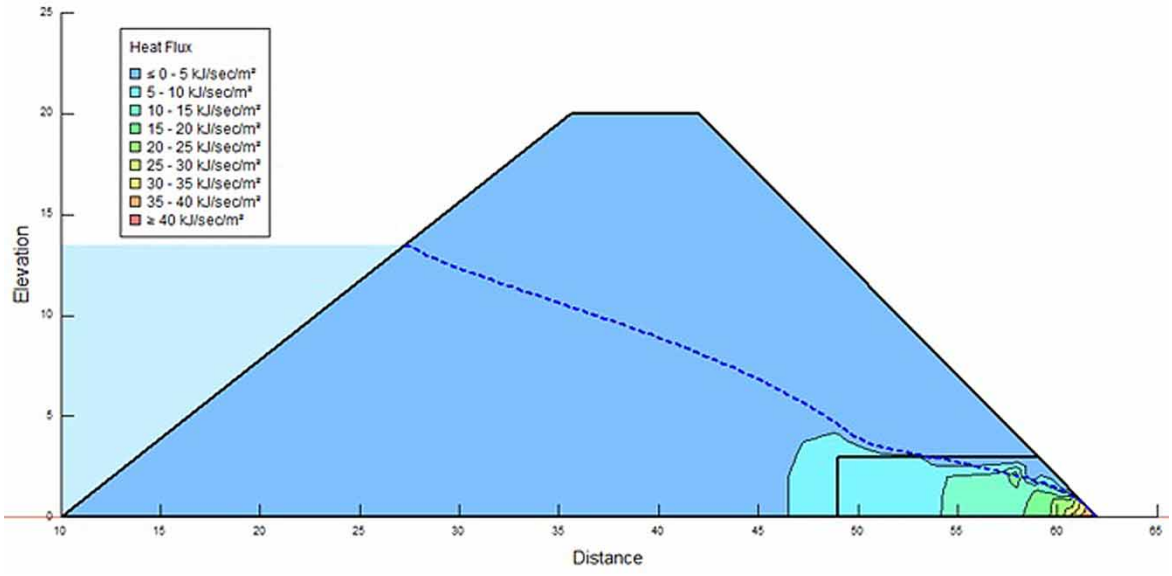


Figure 34 | Heat flux in model 16.

7. CONCLUSION

Sixteen earth dam models constructed in a hydraulic flume were built to study the water flux due to seepage and heat flux due to the convective heat transfer in the respective earth dam models. Temperature measurements were done in the earth dam model placed in a hydraulic flume using digital thermometers to determine the heat flux and validate it with the simulation result of Temp/w in Geostudio 2020. Thus it concluded by observed near walls of a hydraulic flume, soil pores of the dam were of larger volume, resulting in an increased local water flow. Regional flow variation due to wall channeling led to enhanced convection heat transfer, hence increasing the temperature. By increasing the earth dam's length, the heat and water flux had an insignificantly low reduction, and hence flux remained almost the same. With the increase in the dam's height with the same upstream water head, the heat and water flux remained unchanged. Heat flux obtained was found to be a helpful alternative in detecting seepage and water flux in an earth dam. As temperature sensors are cheaper than the

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

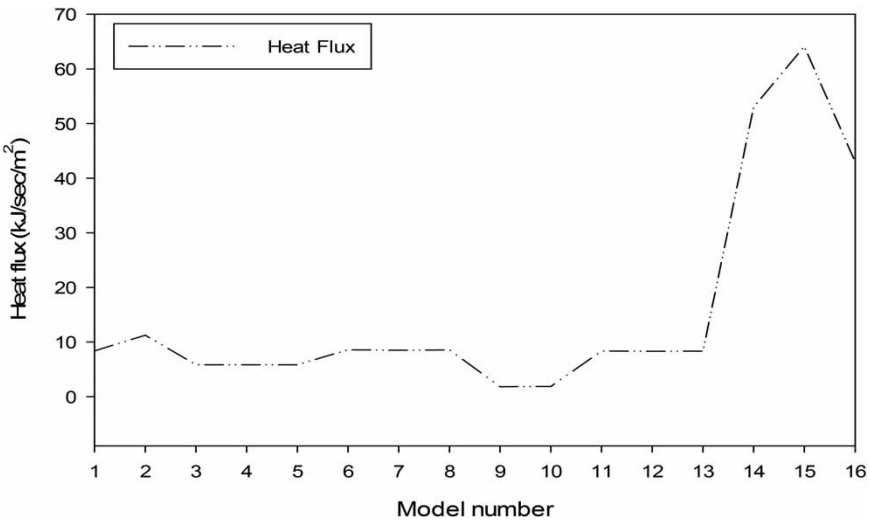


Figure 35 | Heat flux in different models of an earth dam.

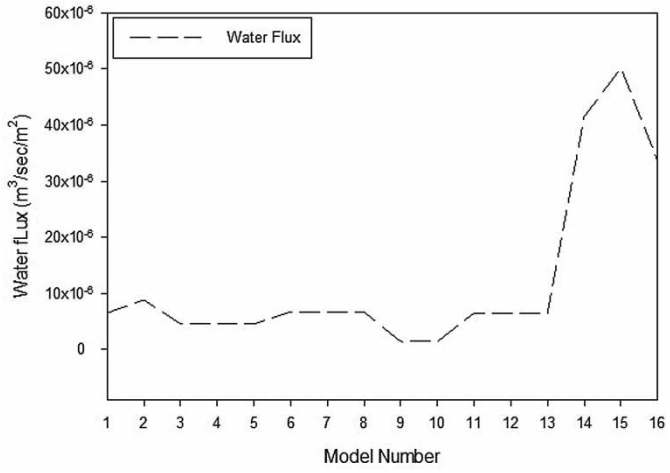


Figure 36 | Water flux in different models of an earth dam.

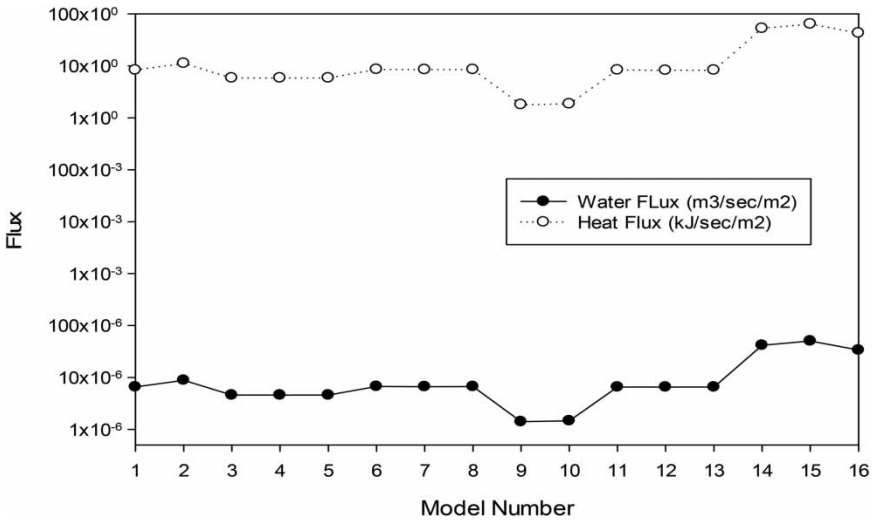


Figure 37 | Water and heat flux comparison in different models of an earth dam.

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

Corrected Proof

discharge and water flow measuring devices, it helps to detect water at a reduced cost compared to conventional devices and hence improves the benefit-cost ratio of the project.

ACKNOWLEDGEMENTS

This work is supported by the Delhi Technological University India. Moreover, Sandeep Mishra's help for the XRD analysis and Lab head, Physics Department of the Delhi Technological University, India, is for the viscosity test.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

REFERENCES

Abu-Hamdeh, N. 2003 Thermal properties of soils as affected by density and water content. *Journal of Biosystems Engineering* **86** (1), 97–102.

Abu-Hamdeh, N. 2014 Specific heat and volumetric heat capacity of granular materials as affected by moisture and density. *Journal of Applied Mechanics and Material* **575**, 103–107.

Alekseevich, A. N. & Sergeevich, A. A. 2017 Numerical modelling on tailings dam thermal-seepage regime considering phase transitions. *Journal of Modelling and Simulation in Engineering* **2017**, 7245413.

Alim, M. A., Ahmed, F. & Islam, M. S. 2017 Seepage analysis of Mahananda earthen embankment at Chapai Nawabganj in Bangladesh. *Journal of American Journal of Engineering and Technology Management* **2** (1), 1–6.

Amanifard, N., Borji, M. & Haghi, A. K. 2007 Heat transfer in porous media. *Journal of Brazilian Journal of Chemical Engineering* **24** (2), 223–232.

Badruddin, I. A., Azeem, Yunus Khan, T. M. & Baig, M. A. A. 2020 Heat transfer in porous media: a mini Review. *Materials Today: Proceedings* **24**, 1318–1321.

Buntebarth, G. 2020 Thermal properties of sand and mineral flours. *Journal of SN Applied Sciences* **2**, 396.

Chen, Y., Zhou, C. & Jing, L. 2009 Modeling coupled THM processes of geological porous media with multiphase flow: theory and validation against laboratory and field scale experiments. *Journal of Computers and Geotechnics* **36**, 1308–1329.

Chuvilin, E. & Bukhanov, B. 2019 Thermal conductivity of frozen sediments containing self-preserved pore gas hydrates at atmospheric pressure: an experimental study. *Journal of Geosciences* **9**, 65.

Cuong, B. Q., Thai, N. C., Yihong, Z. & Chunju, Z. 2017 Estimating seepage in embankment dams based on temperature measurement: a review paper. *Journal of International Journal of Engineering Research and Technology (IJERT)* **6** (1), 106–113.

Del Piero, G. 2020 A mechanical model for heat conduction. *Journal of Continuum Mechanics and Thermodynamics* **32**, 1159–1172.

Engineering ToolBox 2003 *Specific Heat of Some Common Substances*. Available from: https://www.engineeringtoolbox.com/specific-heat-capacity-d_391.html [1/7/2021].

Engineering ToolBox 2004 *Water- Specific Heat*. Available from: https://www.engineeringtoolbox.com/specific-heat-capacity-water-d_660.html [10/28/2020].

Ižvolt, L. & Dobeš, P. 2014 Test procedure for the value of specific heat capacity and thermal conductivity coefficient, XXIII R-S-P seminar, Theoretical Foundation of Civil Engineering (23 RSP) (TFoCE 2014). *Journal of Procedia Engineering* **91**, 453–458.

Johansson, S. & Sjödahl, P. 2004 Downstream seepage detection using temperature measurements and visual inspection monitoring experience from Røsvatn field test dam and large embankment dams in Sweden. In *Proc. Intl. Seminar on Stability and Breaching of Embankment Dams*, p. 21.

Kodešová, R., Vlasáková, M., Fér, M., Teplá, D., Jakšík, O., Neuberger, P. & Adamovský, R. 2013 Thermal properties of representative soils of the Czech Republic. *Journal of Soil and Water Research* **8** (4), 141–150.

Kosky, P., Balmer, R., Keat, W. & Wise, G. 2012 *Exploring Engineering*, 3rd edn. Elsevier, New York, NY, pp. 259–264.

Kurz, D., Alfaro, M. & Graham, J. 2017 Thermal conductivities of frozen and unfrozen soils at three project sites in northern Manitoba. *Journal of Cold Regions Science and Technology* **140**, 30–38.

Misra, A., Becker, B. R. & Fricke, B. A. 1995 A theoretical model of the thermal conductivity of idealized soil. *HVAC&R Research* **1** (1), 81–96.

Radzicki, K. & Bonelli, S. 2010 Thermal Seepage monitoring in the earth dams with impulse response function analysis model. In: *8th ICOLD European Club Symposium*, Innsbruck, Austria, pp. 624–629.

Tokoro, T., Ishikawa, T., Shirai, S. & Nakamura, T. 2016 Estimation methods for thermal conductivity of sandy soil with electrical characteristics. *Journal of Soil and Foundation, The Japanese Geotechnical Society* **56** (5), 927–936.

Yousefi, S., Ghiassi, R., Noorzad, A., Ghaemian, M. & Kharaghani, S. 2013 Application of temperature simulation for seepage inspection in earth-fill dams. *Journal of Gradevinar* **65** (9), 825–832.

First received 8 April 2021; accepted in revised form 13 June 2021. Available online 23 June 2021

Downloaded from http://iwaponline.com/wst/article-pdf/doi/10.2166/wst.2021.241/908469/wst2021241.pdf
by DELHI TECHNOLOGICAL UNIVERSITY (DTU) user

LINGUISTICA ANTVERPIENSIA, 2021 Issue-2
www.hivt.be
ISSN: 0304-2294

High Performance Multipliers Using Hybrid Adder

Sumit Raj[1], Utkarsh Chaurasia[2], Aayush Bahukhandi[3], Poornima Mittal[4]

[1-4] Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India
Email list:smt.raj960@gmail.com, shikaruc@gmail.com, aayushbahukhandi001@gmail.com, poornimamittal@dtu.ac.in

Issue Details
Issue Title: Issue 2
Received: 25 March, 2021
Accepted: 27 April, 2021
Published: 15 May, 2021
Pages: 1547 – 1556

Copyright © 2021 by author(s) and Linguistica Antverpiensia

Abstract

Multiplier is a key component in most of the digital signal processing applications. With the development in the field of VLSI, designers are greatly concerned about achieving low power consumption and high speed. As the multiplier block requires a significant amount of power and plays a significant role in the circuit's speed, optimising it would enhance the circuit's performance. In this paper, a hybrid adder using a combination of transmission gate and pass transistor logic is proposed. Different multipliers are implemented using the proposed hybrid adder. Multiplier with proposed hybrid adder is compared with the multipliers using existing full adder architecture. Performance analysis is done based on different parameters such as delay, power and power-delay product. Power-delay product in multipliers with proposed adder is significantly reduced as compared to multipliers based on mirror and transmission gate adders, implying that the proposed design shows better performance. And in the proposed design, Wallace Tree multiplier and Vedic multiplier show best performance as they have lowest power-delay product, which is approximately half as compared to Bit Array and Baugh Wooley multiplier. All the circuits are designed and simulated using 90nm technology library in the Cadence Virtuoso tool.

Keywords

Hybrid Adder, Array Multiplier, Baugh Wooley Multiplier, Vedic Multiplier, Wallace Tree Multiplier, Power-Delay Product.

1. Introduction

For a chip designer, power consumption and delay are key design factors. Therefore, there was a lot of research undertaken to improve the speed and latency of the multiplier in order to minimize the delay of the overall circuit. With the ever-increasing complication of the processors and the number of transistors in a chip, the power utilization of the system also increases.The addition and multiplier arithmetic devices are key elements of a logic circuit [1]. Full adder is the building block for circuits used for arithmetic operations, such as multiplication, thereby decreasing delay and power in adders, reduces total power and propagation delay of entire circuitry.

The rudimentary mathematical operation for DSP algorithms is multiplication. In signal processing and various other uses, multipliers play a significant role. VLSI architecture for high performance circuits [2] require low power and high-speed multipliers. The multiplicand with the times specified by the multiplier must be repeatedly added to obtain the output. Each bit of the multiplier from the LSB bit is multiplied by the multiplier [3]. Partial products are produced in this method. In order to obtain the end result or product, partial products are to be added.

In this paper, we have attempted to optimize delay and power of multiplier. We have used the approach to firstly improve the performance of full adders by designing a hybrid adder using a combination of TG and PTL based adders. The approach is based on use of low power, minimal transistor count in adders that determine the performance of the mul-

tiplier.We have examined the power and delay characteristics of three different adder circuits which are Mirror Adder(MA), Transmission Gate(TG) based adder and proposed hybrid adder. For the study, we realized the proposed adder which has better performance. By using these adders, four different multipliers: Bit Array, Wallace Tree, Vedic Multiplier and Baugh Wooley are realized. The trade-offs between speed and power of these multipliers were compared. The paper is organized into following section: Section II addresses the different types of adders circuits and its comparative study. Section III describes the multipliers used and its architectures. The simulation technique used is outlined in Section IV. Section V deals with the findings of the simulation analysis and the article is summarized in Section VI, along with the final remarks.

2. Adder

Adders are used for basic addition of binary numbers. It acts like the fundamental unit for multipliers and also the power and speed determining factor in the system. Therefore, it is extremely important to optimize the adder carefully. Hybrid Adders incorporate many schemes to get delay/power optimization. This paper proposes hybrid adders using a combination Pass transistor logic and Transmission gate.

2.1. Mirror Adder

One of the usual ways of implementing the transistor level circuit is the traditional mirror adder. Pull Down and Pull up gate networks are not complementary which means that it is not a CMOS logic. This allows some of the transistors to be removed. Mirror Adder uses 24 transistors. The logical expressions for sum(S) and output (C_0) are

$S = AB'Ci' + A'B'Ci + A'BCi' + ABCi$

$C_0 = AB + BCi + ACi.$

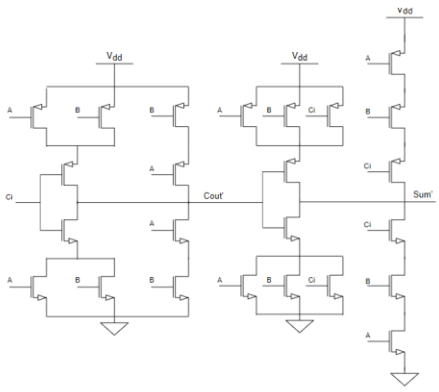


Figure 1. Architecture of Mirror Adder.

2.2. Transmission Gate

A transmission gate comprises of a parallel NMOS and PMOS transistor with their gates constrained by complementary logic [5]. A 0 is passed using NMOS and a 1 is passed using PMOS in a transmission gate. A transmission gate based multiplexer is a better option for decreasing both delay and power as gate level architecture requires intricate hardware. Thus, transmission gate should be and is used to implement a multiplexer [6].
Adder based on transmission gate comprises of three modules:

1. First module contains XOR gate with inputs as X and Y
2. Second module is an XOR gate for sum output.
3. Third module is a two-bit MUX used to generate Cout.

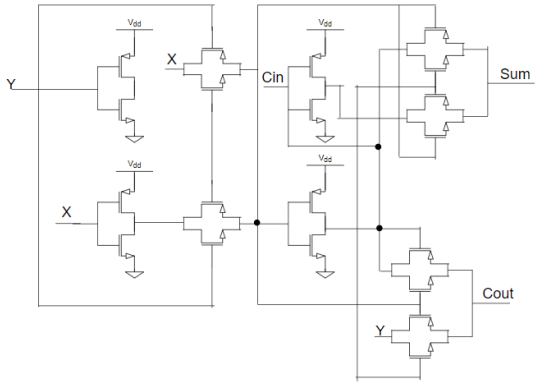


Figure 2. Architecture of Transmission gate-based Adder.

2.3. Proposed Adder

Different full adder like mirror adders (MA) and full adders, based on the transmission gate (TG) and pass transistor logic (PTL) requires a large number of transistors for their design, which increases the area of the circuit. The power consumption and propagation delay depend upon the techniques used in design. A low output voltage swing affects the proposed 14T approximate adder. Due to which, two inverters [15] are employed at the output which acts as level restorer [16]. Propagation delay, power dissipation, and circuit area increases with the use of buffers, but improves the output waveform. The proposed 18T circuit is designed as a mixed logical architecture that utilizes a transmission gate and pass transistor logic combination. In this paper, a hybrid adder using 18-transistor (18T) logic has been implemented. The proposed 18T hybrid adder provides a sturdy output voltage swing than approximate 14T transistor. This hybrid adder is used to implement different multiplier architectures.

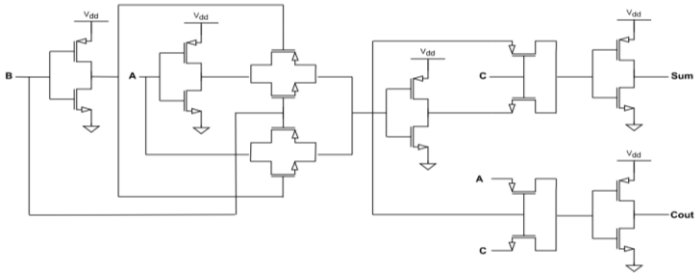


Figure 3. Architecture of Proposed Adder

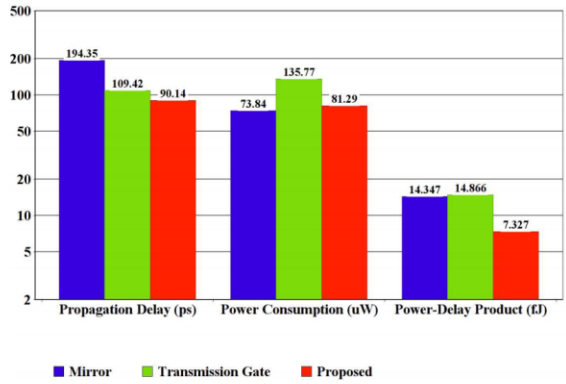


Figure 4. Graphical representation of comparison of different Adders

ISSN: 0304-2294

Types of Adders	Propagation Delay (ps)	Power Consumption (uW)	Power-Delay Product (fJ)
Mirror [4]	194.35	73.84	14.347
Transmission Gate [5]	109.42	135.77	14.866
Proposed	90.14	81.29	7.327

Table I. Comparison of different Adders

3. Multiplier

Multipliers are one of the most important hardware blocks of a lot of DSP systems. Digital filters, digital communications and spectral analysis are popular DSP applications in which the multiplier plays an important role. Many existing DSP implementations are aimed at portable batteries, making power dissipation one of the principal design constraints. Two inputs and one output are used in the multiplier. The multiplier and multiplicand are assigned as two inputs. The multiplicand is repeatedly added to the times specified by the multiplier to get the output. The multiplicand is multiplied by every bit of the multiplier starting from the LSB bit. Any bit of the multiplier starting from the LSB bit multiplies the multiplicand. Partial products are produced in this method. To get the final result, partial products have to be added.

In this paper, we have used partial products. AND gates are used to generate partial products. A 4x4 bit multiplier contains 16 partial products. The cumulative number of AND gates thus used amounts to 16. According to notation Snm means nth bit of A number is an input to AND gate along with mth bit of B, i.e., S_{10} – 1st bit of A AND 0th bit of B.

3.1. Array Multiplier

A Digital combinational circuit that uses a range of half and full adders to get the product of two binary numbers is called an array multiplier [7,14]. With the filter coefficients, the input signals can be multiplied. A multiplier module can do this. An array multiplier comprises of two full adder arrays in two dimensions. Combination multipliers consist of multipliers of the array. The combinational multiplier contains rows of carry save adder circuits. Parallel to AND gates, the summands are measured and then added to the single-bit adder list. A nxn multiplier array consists of (n-1) rows of the carry save adders with a full adder in each row. The last row is a ripple adder intended for carry transmission. The inputs given are x, y, and Cin. The outputs are Sum(S) and are executed (Cout). The rules to execute the multiplication of arrays are:

1. Matrices need to have equal dimensions.
2. The corresponding matrix dimensions would be the same when two matrices are multiplied.
3. The resulting elements are the product of the matrices multiplied by the corresponding elements. The multiplication steps can be seen in the figure.

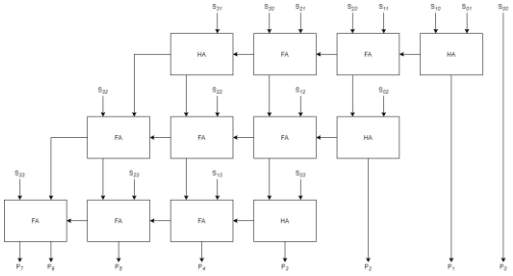


Figure 5. Architecture of Array Multiplier

3.2. Baugh Wooley Multiplier

Charles R. Baugh and Bruce A. Wooley, proposed "A Two's Parallel Array Multiplication Algorithm." The parallel array multiplication is represented in m-bit by n-bit for this.

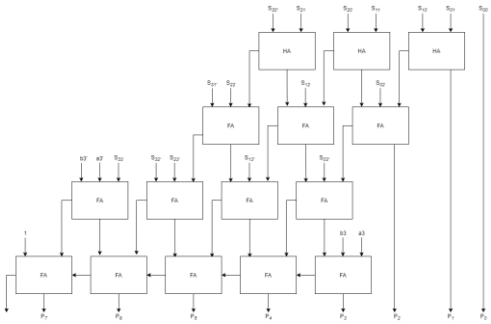


Figure 6. Architecture of Baugh Wooley Multiplier.

Due to the need to operate digital filters and Fourier transform processors, multiplication is an integral feature of the digital system.The Baugh-Wooley [8] series multiplier is an effective way to multiply signed and unsigned numbers. In the High-Performance Multiplier (HPM) tree, the Baugh Wooley Algorithm is used to inherit the standard array multiplier and recurring structure. The multiplier Baugh Wooley shows less delay, a low power dissipation and occupies a lesser area compared with other multipliers. Baugh Wooley's design is built on the carry save algorithm.

3.3. Vedic Multiplier

Vedic Mathematics system is composed of 16 sutras. Urdhava Tiryakbhyam (Vedic Multiplier) [9,10] means vertical and cross-wire in English. Coherence is the important aspect of the Vedic system. Using Vedic System, even difficult and challenging problems can be solved easily as compared to Modern Mathematics. Thus, Vedic System seems more structured and systematic than the Modern System.In signal processing techniques, the most significant operation is multiplication. The main hurdle that slows down the system is the number of computational steps involved and the delay associated with each step. Thus, the speed of the multiplier determines the efficiency of system. The performance of the system is mainly decided by three factors, i.e., power, delay and power delay product.

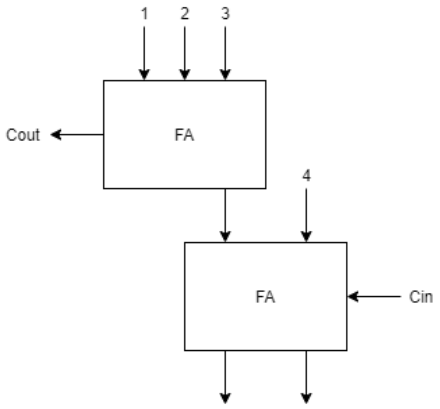


Figure 7. 4:2 Compressor

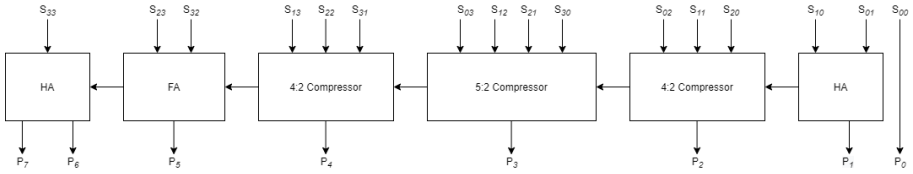


Figure 8. Architecture of Vedic Multiplier

3.4. Wallace Tree Multiplier

Wallace tree design [11,12] is an optimal approach to the deployment of hardware to minimize calculation time when you use carry save adders to add multiple operands. The research work of othersalso suggest that speed of Wallace Tree multipliers is higher than array multiplier[13].It implements simple and effective hardware methodology, using column compression methods to multiply integers. Wallace tree provides quick speed since the overall delay is proportional to the logarithm of the length of operand of the multiplier instead of linear dependence, as is the case with the array multiplier. The Wallace Tree Multiplier operation includes three stages: partial product formation, grouping of such partial products and addition of those products using adders.

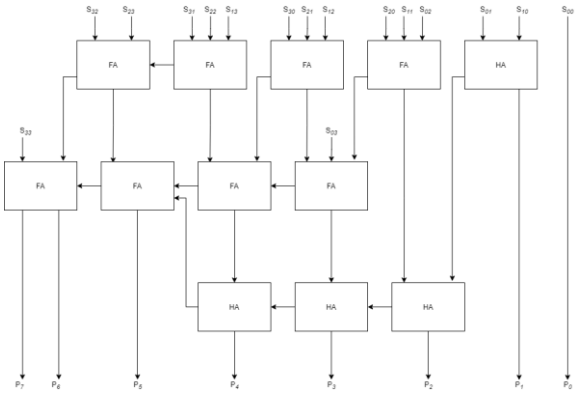


Figure 9. Architecture of Wallace Tree Multiplier

4. Simulation and Results

A comprehensive comparison of all the four multipliers is presented using different adders like mirror adder, transmission gate-based adder and proposed hybrid adder. For analysis purpose, different parameters were considered namely power consumption, propagation delay and power-delay product. The results were obtained using Cadence Virtuoso GPDK 90nm library.

4.1. Propagation Delay

Table II and Figure shows that propagation delay is significantly reduced in multipliers with proposed hybrid adder. As compared to mirror adder and transmission gate-based adder, the propagation delay using proposed adder is 31.06% and 23.98% less in Wallace Tree multiplier, 11.82% and 14.81% less in Bit Array multiplier, 44.79% and 33.42% less in Baugh Wooley multiplier, and 45.44% and 35.43% less in Vedic multiplier. The Bit Array multiplier has the highest delay, while the Vedic multiplier has the lowest in the proposed design.

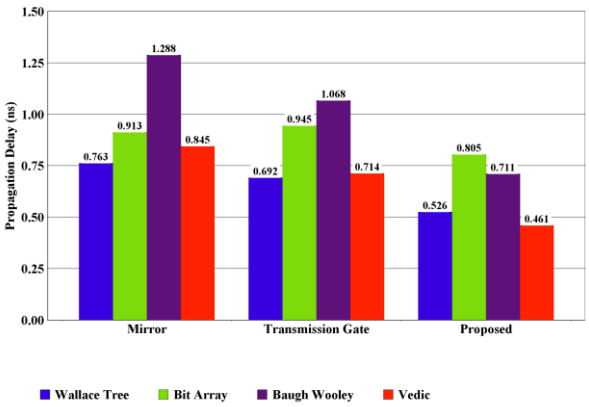


Figure 10. Graphical Representation of Comparison of Propagation Delay

4.2. Power Consumption

From the Table I, II and Figure, it can be concluded that the power consumption of the proposed adder is much lower than that of the TG based adder and is equal to that of the mirror-based adder. As a result, the power dissipation of the multiplier using the proposed adder is lower than that of the TG adder-based multiplier and is equivalent to that of the mirror adder-based multiplier. The power consumption using proposed adder is highest in Bit Array multiplier, but it is reduced slightly in Baugh Wooley multiplier (by 13.75%), however it is dramatically reduced in Vedic multiplier (by 49.6%), and then further reduced (by 15%) in Wallace Tree multiplier.

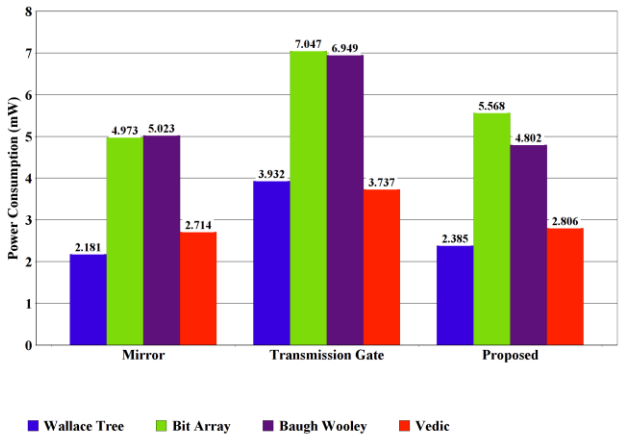


Figure 11. Graphical Representation of Comparison of Power Consumption

4.3. Power-Delay Product

Performance of a circuit design is determined by the parameter Power-Delay Product. And circuit having lower PDP shows a better performance. Improved performance i.e., a lower power-delay product, is only feasible when the circuit's power-delay optimizations are well realized. The PDP of the proposed multipliers is much lower than that of multipliers based on mirror and TG adders, implying that the proposed design performs much better.

Table II and Figure shows that PDP in Baugh Wooley multiplier (proposed) and Vedic multiplier (proposed) is approximately half that of mirror-adder and TG adder-based multipliers. And in the proposed design, Wallace Tree multiplier and Vedic multiplier show better performance as they have lowest PDP, which is almost half the value as compared to Bit Array multiplier (reduction around 70 %) and Baugh Wooley multiplier (reduction around 60 %).

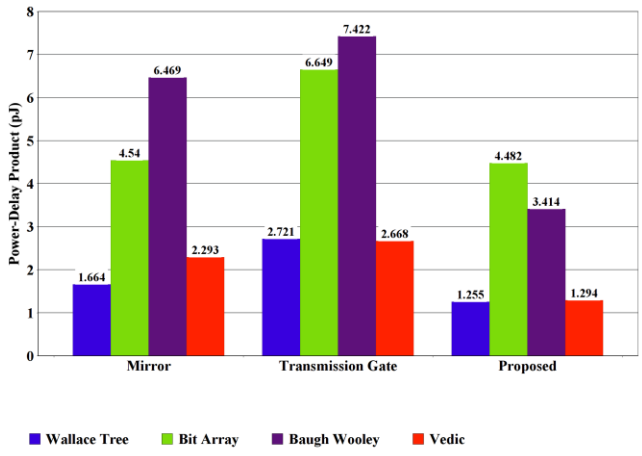


Figure 12. Graphical Representation of Comparison of Power-Delay Product

Types of Adders	Types of Multipliers	Propagation Delay (ns)	Power Consumption (mW)	Power-Delay Product (pJ)
Mirror	Wallace Tree	0.763	2.181	1.664
	Bit Array	0.913	4.973	4.540
	Baugh Wooley	1.288	5.023	6.469
	Vedic	0.845	2.714	2.293
Transmission Gate	Wallace Tree	0.692	3.932	2.721
	Bit Array	0.945	7.047	6.649
	Baugh Wooley	1.068	6.949	7.422
	Vedic	0.714	3.737	2.668
Proposed	Wallace Tree	0.526	2.385	1.255
	Bit Array	0.805	5.568	4.482
	Baugh Wooley	0.711	4.802	3.414
	Vedic	0.461	2.806	1.294

Table II. Comparison of different Multipliers

Conclusion

In this paper,a hybrid adder using a combination of transmission gate and pass transistor logic is proposed and different multipliers are implemented using the proposed hybrid adder. A comprehensive evaluation of various parameters, such as power dissipation, propagation delay and Power-Delay product, has been presented using extensive simulation results. The multipliers that use the proposed adder in its logic circuit shows more optimal values in the case of each power dissipation, propagation delay and Power-Delay product.Power-Delay product is substantially reduced for multipliers that have a proposed adder compared with mirror- and transmission-gate-adders which implies improved performance in the proposed configuration.The proposed design shows best results for Wallace Tree multiplier as well as Vedic multiplier which have the lowest power-delay product that is about half as the Bit Array and Baugh Wooley multipliers.

References

[1] B. Kumar, K. Raj and P. Mittal, “FPGA implementation and mask level CMOS layout design of redundant binary signed digit comparator, *International Journal of Computer Science and Network Security*, Vol. 9,Sept 2009, pp. 107-115.

[2] P. Mittal and N. Kumar, "Comparative analysis of 90nm MOSFET and 18nm fin-FET based different Multiplexers for Low Power Digital Circuits", *International Journal of Advanced Science and Technology*, Vol. 29,2020, pp.4089-4096.

[3] A.P.S. Rathod, P. Lakhera, A.K. Baliga P. Mittal and B. Kumar, "Performance Comparison of Pass Transistor and CMOS Logic Configuration based De-Multiplexers,"*International Conference on Computing Communication and Automation* (ICCCA- 2015), May 2015, pp. 1433-1437.

[4] K. M. Reddy, Y. B. N. Kumar, D. Sharma and M. H. Vasantha, "Low power, high speed error tolerant multiplier using approximate adders," 2015 *19th International*

Symposium on VLSI Design and Test, 2015, pp. 1-6.

[5] Z. Yang, J. Han and F. Lombardi, "Transmission gate-based approximate adders for inexact computing," *Proceedings of the 2015 IEEE/ACM International Symposium on Nanoscale Architectures*, 2015, pp. 145-150.

[6] H. Jiang, J. Han, & F. Lombardi, "A comparative review and evaluation of approximate adders" *25th edition on Great Lakes Symposium on VLSI* , pp. 343-348.

[7] S. Krishna, B. Vamsi, and K. Dhanunjaya. "Analysis and modelling of low power array multipliers using cadence virtuoso simulator 45 nm technology." *International Journal of Electronics, Communication & Instrumentation Engineering Research and Development (IJECIERD)* Vol. 3, Issue 4, Oct 2013, 49-64.

[8] P. Karuppusamy, "Design and Analysis of Low-power. High-Speed Baugh Wooley Multiplier." *Journal of Electronics 1.02* (2019), pp. 60-70.

[9] Pushpavalli, M., C. Ragapriya, and B. Praveena. "Compressors Based High Speed 8 Bit Multipliers Using Urdhava Tiryakbhyam Method." *International Journal of Engineering and Management Research (IJEMR)* 7.1 (2017), pp.127-131.

[10] G. Rawat, K. Rathod, S. Goyal, S. Kala, P. Mittal, "Design and Analysis of ALU: Vedic Mathematics Approach*," International Conference on Computing Communication and Automation (ICCCA- 2015)*, May 2015, pp. 1372-1376.

[11] A. Gopal, and M. Naresh. "Comparative Analysis of 11T and 16T and 28T Full Adder Based 4* 4 Wallace Tree Multiplier Using Cadence 180nm Technology.", *International Research Journal of Engineering and Technology (IRJET)*, Volume: 04 Issue: 08, Aug -2017.

[12] I. Hussain, and M. Kumar. "Design and Analysis of a Conventional Wallace Multiplier in 180nm CMOS Technology." *IOSR Journal of VLSI and Signal Processing* 5.1 (2015), pp.60-65.

[13] I. Hussain, R. K. Sah, and M. Kumar. "Performance Analysis Comparison of a Conventional Wallace Multiplier and a Reduced Complexity Wallace multiplier." *IOSR Journal of VLSI and Signal Processing* 5.2 (2015), pp. 23-27.

[14] S. Karunakaran, "Exploration on Power Delay Product of various VLSI Multiplier Architectures. *"International Journal of Mechanical Engineering and Technology (IJMET)* Volume 9, Issue 1, January 2018, pp. 53–59.

[15] P. Mittal, "Ditch and Elevated OTFT based Improved Common Source Voltage Amplifier: Frequency Response Characteristics and Analytical Modeling", *Journal of the Society for Information Display*, 2021, pp. 1-17.

[16] N. Mishra, P. Mittal and B. Kumar, "Analytical Modelling for Static and Dynamic Response of Organic Pseudo All-p Inverter Circuits", *Journal of Computational Electronics*, Vol. 18, Issue 4,Oct – 2019, 1490–1500.

arXiv:2105.14515v1 [cs.CL] 30 May 2021

How Low is Too Low?
A Computational Perspective on Extremely Low-Resource Languages

Rachit Bansal[1] **Himanshu Choudhary**[1] **Ravneet Punia**[1]
Niko Schenk[2†] **Jacob L Dahl**[3] **Émilie Pagé-Perron**[3]
[1] Delhi Technological University [2] Amazon Berlin, Germany [3] University of Oxford
{rachitbansal2500, himanshu.dce12, ravneet.dtu}@gmail.com
nikosch@amazon.com
{jacob.dahl, emilie.page-perron}@wolfson.ox.ac.uk

Abstract

Despite the recent advancements of attention-based deep learning architectures across a majority of Natural Language Processing tasks, their application remains limited in a low-resource setting because of a lack of pre-trained models for such languages. In this study, we make the first attempt to investigate the challenges of adapting these techniques for an extremely low-resource language – Sumerian cuneiform – one of the world's oldest written languages attested from at least the beginning of the 3rd millennium BC. Specifically, we introduce the first cross-lingual information extraction pipeline for Sumerian, which includes part-of-speech tagging, named entity recognition, and machine translation. We further curate *InterpretLR*, an interpretability toolkit for low-resource NLP, and use it alongside human attributions to make sense of the models. We emphasize on human evaluations to gauge all our techniques. Notably, most components of our pipeline can be generalised to any other language to obtain an interpretable execution of the techniques, especially in a low-resource setting. We publicly release all software, model checkpoints, and a novel dataset with domain-specific preprocessing to promote further research.

1 Introduction

Sumerian is one of the oldest written languages, attested in the cuneiform texts from around 2900 BC and possibly the language of even older proto-cuneiform texts from the second half of 4th millennium BC (Englund, 2009). Specialists in Assyriology have recently worked to digitize Sumerian scripts, annotate, and translate a part of them to modern-day languages like English and German.



obverse.
1. 1(disz) kusz udu niga
1 hide, grain-fed sheep;

2. 1(disz) kusz masz2 niga
1 hide, grain-fed goat;

3. kusz udu sa2-du11
sheep hides, regular offerings,

4. ki {d}iszkur-illat-ta
from Adda-illat,

reverse.
1. a-na-ah-i3-li2
Anah-ili;

2. szu ba-an-ti
did receive.

3. iti ezem-an-na
Month: An-festival,

4. mu na-ru2-a-mah mu-ne-du3
Year: He erected the great stele for them.

Figure 1: Tablets inscribed with Sumerian cuneiform script, their corresponding digitized transliterations, and human-translated English text for each line.

In this work, we attempt to create the first information extraction and translation pipeline for Sumerian. Specifically, we focus on machine translation from Sumerian to English, and sequence labeling tasks of Named Entity Recognition (NER) and Part of Speech Tagging (POS).

Figure 1 shows a sample of our raw data where the Sumerian text has been derived from the tablet-inscribed cuneiform script along with its human-interpreted English translations. Creating an annotated corpus for such a language is a tedious task. Thankfully, we obtain our data from openly available sources and corpora, painstakingly annotated and translated by human experts. Yet, for languages like Sumerian, which are not fully-understood by humans themselves, transferring knowledge and patterns to learning algorithms from this limited data becomes extremely difficult. The consequent challenge posed for NER and POS is evident. Lack of annotated data and fuzzy character-level text

Datasets and training subroutines are available at linktr.ee/rachitbansal

[†]Work was done prior to joining Amazon at Goethe University Frankfurt

makes it hard for a model to generalise, irrespective of its size.

In case of machine translation, the labeled data is composed of incomplete and short phrase-like sentences, specially on the target-side. This makes the context largely ambiguous. Moreover, we find that for a majority of medieval languages the target-side translated text is highly incoherent with modern-day English language text, making it impossible to use the latter in semi-supervised and unsupervised settings.

Throughout this study, we elaborate on such challenges induced when working with low-resource languages, and talk about what makes some of these languages like Sumerian 'extremely' low-resource. Through extensive experimentation, evaluation, and analysis we further introduce specific algorithms and modifications to work around them.

In all, our contribution is three-fold:

1. Building and analyzing a variety of algorithms on the unexplored human-annotated Sumerian dataset for sequence labeling tasks of POS Tagging and NER. (§3)
2. Introducing the problem of *Target-side Incoherence* for low-resource settings and its effect on semi-supervised and unsupervised machine translation (§4.2). Further investigating specific modifications and methodologies to cope-up with these constraints. (§4)
3. Introducing *InterpretLR*, a generalisable toolkit to interpret low-resource NLP. We apply to additionally study, compare, and evaluate all of the proposed techniques for machine translation and sequence labeling. (§7)

Throughout this work, we have conducted human studies and evaluation for our models, in addition to automated metrics. For studying our models with *InterpretLR*, we have made use of human annotations.

2 Background

2.1 Data

Sumerian is an ancient language from Iraq that was written using the cuneiform script. While Basque and Turkish display some similarities (split-ergativity, agglutinativity), it is a language isolate (Englund, 2009). We have found artifacts dating to around 2900 BC with Sumerian texts inscribed until the first century AD. Most of the Sumerian texts found to this day are administrative in nature as, during the third dynasty of the Ur III Period, the state administration swell to an unprecedented level of activity which was not seen again later in the history of Mesopotamian culture. All through this study, our evaluation sets are composed of Ur III Admin text only and it acts as our in-domain data.

Part of the datasets we used were assembled from the Cuneiform Digital Library Initiative (CDLI)[1], Machine Translation and Automated Analysis of Cuneiform languages (MTAAC) project (Pagé-Perron et al., 2017)[2] and The Electronic Text Corpus of Sumerian Literature (ETCSL) dataset [3]. CDLI and MTAAC datasets contain the Ur III Administrative (Admin) texts[4] which are preserved by the CDLI[5]. The MTAAC and ETCSL corpora were both manually annotated for morphology by cuneiform linguistics.

We divided the data between training and testing sets, and then to reduce the data sparsity, we performed text augmentation using a set of labeled named entities for these sets separately. This increased our combined number of phrases from $25{,}000$ to $48{,}000$, representing our final dataset for sequence labeling. Figures 2 and 3 provide the distribution of word tokens in our final pre-annotated dataset. The corpus consists of phrases with lengths ranging from 1 to 19 words. These phrases are small since they are translated line by line from the scripts. Around $2{,}500$ phrases were used for testing, while the $45{,}500$ were employed for training purposes.

For machine translation, the final dataset summarizes as (i) $10{,}520$ parallel phrases from the Ur III administrative corpus; (ii) $88{,}460$ parallel phrases, all genres combined; and (iii) all monolingual Sumerian data (1.43 million phrases). In all cases, phrases are short, generally ranging from 1 to 5-word tokens.

2.2 Related Work

Past work aimed at machine translation of Sumerian-English (Pagé-Perron et al., 2017; Punia et al., 2020) have used the minimal bitext upon a variety of general statistical and neural supervised

[1] https://cdli.ucla.edu

[2] https://cdli-gh.github.io/mtaac/

[3] http://http://etcsl.orinst.ox.ac.uk/

[4] The Third Dynasty of Ur is a cultural and temporal period ranging in $\sim 2112 - 2004$ BC, in Mesopotamia

[5] https://github.com/cdli-gh/data, https://github.com/cdli-gh/mtaac_gold_corpus/tree/workflow/morph/to_dict

techniques. However, they do not handle the text-level peculiarities any differently than one would do for a high-resource language, thus, often failing to capture context, resulting in poor and inconsistent translations. Techniques, learning algorithms, and architectures that optimally use the vast monolingual data and parallel sentences while keeping in mind the several linguistic limitations are motivated in such a scenario. Thus, we experiment on semi-supervised and unsupervised techniques across the three categories of Data Augmentation (Sennrich et al., 2016; He et al., 2016), Knowledge Transfer (Zoph et al., 2016), and Pre-training (Conneau and Lample, 2019; Song et al., 2019).

In the past, Pagé-Perron et al. (2017) applied statistical models for morphological analysis and information extraction for Sumerian. Although, due to the unavailability of annotated data, these models could not generalise well. Liu et al. (2015) and Luo et al. (2015) used an unsupervised approach for NER with the help of domain experts and used Contextual and Spelling rules to build the model. They also post-processed their outputs automatically, which enhanced their results. In this work, we thoroughly investigate a wide range of algorithms for these sequence labeling tasks and consequently take a first step towards effective information extraction for Sumerian.

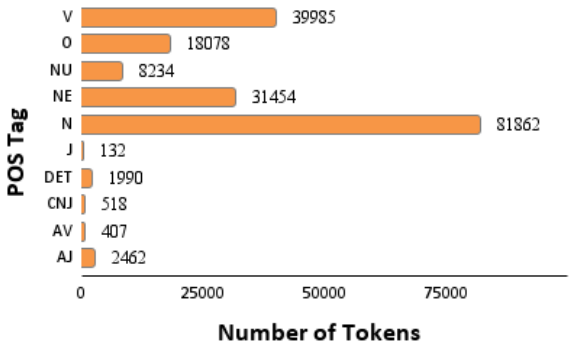


Figure 2: Composition of the POS tagging dataset. Here, "NE" stands for named entities, "O" stands for unstructured words. Other tags are in accordance with ORACC.

3 Part of Speech Tagging and Named Entity Recognition

In this section, we talk about the various algorithms that we investigated to carry out the sequence labeling tasks of POS and NER for Sumerian. The subsequent experimental results are described and discussed in Section 6.

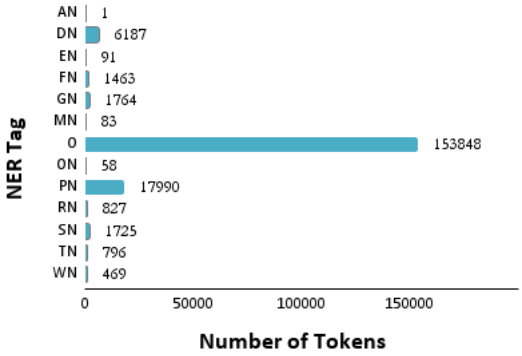


Figure 3: Composition of our NER dataset. Tags are in accordance with ORACC.

Conditional Random Fields CRF (Lafferty et al., 2001) is a discriminative probabilistic classifier, which optimises the weights or parameters in order to maximize the conditional probability distribution $P(y \mid x)$. They take set of input features (language or domain specific) into account, using the learned weights associated with these features and previous labels to predict the current label. Since CRFs use feature sets (rules) which are language-specific, it makes the model more robust specially for very low-resource languages. In our case we developed domain specific rules with the help of previous studies (Liu et al., 2015; Luo et al., 2015) and language experts. A set of these rules are mentioned in the Appendix.

Bi-directional LSTM We also experiment across Recurrent Neural Networks (RNNs) to deal with the sequential text input. We employ Bi-LSTM (Hochreiter and Schmidhuber, 1997; Schuster and Paliwal, 1997) in particular. As in Huang et al. (2015), an additional CRF layer is used for efficient usage of sentence level tag information and past input features by LSTM cells.

FLAIR Akbik et al. (2018) introduced a Contextual String Embedding for Sequence Labeling, FLAIR, which has shown great success for various languages for NER (Akbik et al., 2019b). We make use of the two distinct properties of its embeddings: (i) training without any explicit notion of words and fundamentally modeling the words as a sequence of characters, and (ii) deriving and using the context from surrounding tokens.
We train the bi-directional character language model using the Sumerian monolingual phrases and retrieve the contextual embedding for each word which we then pass into the vanilla Bi-LSTM CRF model.

RoBERTa We also investigate the transformer-based language model, RoBERTa (Liu et al., 2019). The encoder is first pre-trained on our Sumerian monolingual data, and then fine-tuned on our downstream sequence labeling tasks using the labeled data.

4 Machine Translation

In this section, we present our experiments for machine translation, primarily focusing on specific data and algorithmic modeling techniques which may be generalised for any extremely low-resource language that may or may not suffer from *Target-side Incoherence*, a phenomenon which we also introduce herein. All results are summarised in Table 1.

4.1 Supervised NMT

In order to create a benchmark for the semi-supervised and unsupervised approaches, we perform supervised machine translation using the limited bitext available ($\sim$10,000 phrases). We perform experiments on a variety of data configurations which are given by:

1. `UrIIISeg`: Follows the format as present in the original texts provided by Assyriologists and used in the past attempts for Sumerian-English machine translation (Pagé-Perron et al., 2017). It contains only the in-domain Ur III Admin Data with line-by-line translated segments of 1-5 words each, amounting to 10528 segments.
2. `UrIIIComp`: Also contains the in-domain data only, but multiple segments are concatenated together to form complete sentences. The 'completeness' of a sentence is ensured using punctuation marks. It comprises of only 4792 sentences.
3. `AllSeg`: Contains out-of-domain Sumerian text segments in addition to in-domain Ur III Admin alone. The additional text varies across a wide range of genres such as literary, lexical, ritual and legal, resulting into a corpus size of 88466 segments.
4. `AllComp`: Combines the additional features of 2. and 3., thus comprising of a total of 32694 complete text sentences from all genres.

We make use of the vanilla transformer encoder and decoder architecture (Vaswani et al., 2017) for all our supervised machine translation experiments over these three different bitext configurations. Noting the results as in 1, the `AllComp` text configuration is used for all other experiments. The computational configurations are mentioned in Section 5.

4.2 Semi-Supervised and Unsupervised NMT

We observed that one of the primary reasons for the lack of success of semi-supervised and unsupervised algorithms for low-resource settings, specially for medieval languages, is *the lack of coherence between monolingual texts in the modern-day corpora to the target-side text in the available parallel corpora*. We refer to this as the ***Target-side Incoherence*** (*TSIC*) problem for such languages.

As can be seen from Figure 1, the transliterated English text in our parallel corpora is vastly different from general modern-day English texts. In Sumerian, this is because the text has been human-translated to English on the level of words and small segments due to insufficient knowledge of the language. This results into a contextually distorted English language text, as compared what we see in general corpora. This leads to multiple pitfalls. Most significantly, the colossal monolingual data available for a data-rich target-side language (i.e., English in this case) can no longer be used. This *Target-side Incoherence* holds true for most medieval language texts like Sumerian, which makes them 'extremely' low-resource.

In this section, we elaborate on the problems caused due to *TSIC* and further present our hypothesis on adapting various semi-supervised and unsupervised NMT techniques to deal with them.

Forward Translation Back-translation (BT) (Sennrich et al., 2016) has been widely used and analysed for NMT across a large set of language pairs. BT uses a reverse model, Sumerian $\leftarrow$ English trained on the existing parallel corpora, when the task is to translate from Sumerian $\rightarrow$ English, and applies it on the target-side monolingual corpus. The synthetic samples thus generated are added to the source-side corpus and a new reverse model is trained on the augmented dataset. It has been shown to outperform its forward counterpart, Forward Translation (FT) (Zhang and Zong, 2016; Burlot and Yvon, 2018), which instead uses a forward (Sumerian $\rightarrow$ English) model to augment

the target-side of the bitext.
However, due to *TSIC*, the target-side monolingual data falls into a completely different distribution than what a Sumerian ← English model is trained on. Using back-translation in such a scenario results into a poor source-side augmentation, doing more harm than good. Keeping this in mind, we rely on forward-translation (FT), thus using the Sumerian monolingual text.

We divide the Sumerian monolingual data into 8 shards, each containing ∼100,000 monolingual `AllComp` sentences each. The FT process takes place for each shard and the Transformer model is trained after each shard is forward-translated.

Large scale studies (Edunov et al., 2018; Wu et al., 2019) have shown the heavy dependency of BT and FT on aspects like sampling methods and the amount of parallel data. The performance with non-MAP (where, MAP stands for *maximum a posteriori*) estimation methods like Nuclear Sampling (Holtzman et al., 2018) and Beam Search with Noise improves almost-linearly with the amount of bitext, and thus, for low-resource settings (∼80,000 sentence pairs), MAP methods have been shown to give better results. This was also observed in our experiments and the reported results are obtained using Beam Search (§5).

Cross-Lingual Language Model Pre-training We further make use of XLM (Conneau and Lample, 2019) to carry out a wide range of experiments for both unsupervised and semi-supervised fine-tuning techniques. Considering the lack of original target monolingual text due to *TSIC*, the following target data configurations were used for pre-training the XLM:

1. `WMT`: Ignoring the lack of coherence between general English texts and the evaluation + training texts, to compose the entire target monolingual data with the WMT '18 English Texts. Amounts to a total of 20M sentences.
2. `Orig`: Composed of all the English side texts in `UrIIISeg`, `UrIIIComp`, `AllSeg` and `AllComp` combined. Contains only ∼60,000 sentences.
3. `Mixed`: This consists of all of 2. and as many sentences as 1. through which the net size of the corpus equalizes the Sumerian monolingual, i.e., 1.5M.

In the pre-training phase, we perform various experiments over different combinations of MLM and TLM. It is further fine-tuned on a denoising auto-encoding objective for Unsupervised while cross-reference machine translation objective over the parallel data for semi-supervised training. BT steps are also performed in both cases.

Data Augmentation In order to further reduce the effect of *TSIC* on the model performance and to allow the model to attend to a larger and more diverse volume of target text during pre-training, we make use of the following data augmentation techniques:

1. `BERT`: Replacing words by the spatially closest words measured by Cosine Similarity in BERT (Devlin et al., 2019) Embeddings, with a threshold of 0.8.
2. `WordNet`: Replacing words with WordNet (Miller et al., 1990) synonyms.
3. `CharSwap`: Introduces certain character-level perturbations in the text by substituting, deleting, inserting, and swapping adjacent character tokens.

Different combinations of these techniques have been used to augment the `Orig` type target monolingual data. The resultant target-side corpora sizes are summarised in Figure 4.

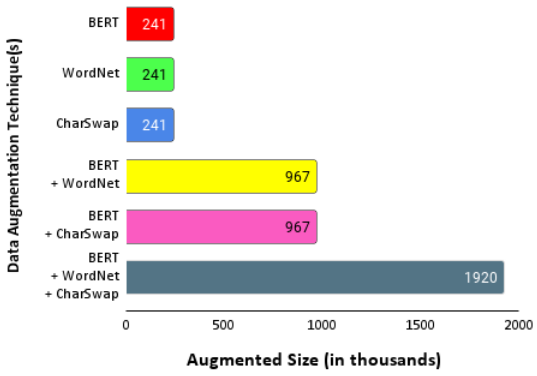


Figure 4: Effective size of the target monolingual corpora with different combinations of augmentation.

5 Experimental Setup

All our experiments have been implemented in PyTorch, except for the Bi-LSTM and CRF which were done in Tensorflow. In addition to this, we used FairSeq (Ott et al., 2019), FLAIR (Akbik et al., 2019a), HuggingFace Transformers (Wolf et al., 2019), and Open-NMT (Klein et al., 2017) frameworks in Python. Nvidia Apex was used for memory optimisation using fp-16 training. Experiments related to Bi-LSTM, CRF, vanilla transformers, and FT were performed on a single $8GB$ Nvidia

Technique	S	US	SS	HE
Vanilla Transformer				
UrIIISeg	36.32			2.202
UrIIIComp	33.45			2.242
AllSeg	37.01			2.360
AllComp	42.23			2.431
+3×FT*			41.98	2.358
+5×FT			**44.14**	**2.504**
+7×FT			42.95	2.367
XLM				
MLM, Orig		4.49	15.04	
MLM + TLM, WMT		0.94	–	
Mixed		13.08	21.23	1.104, –
Orig		12.73	24.64	1.294, –
XLM + Data Augmentation				
BERT		13.06	29.50	1.320, 1.704
WordNet		13.08	28.57	1.269, 1.690
CharSwap		12.92	29.04	
BERT+WordNet		13.34	26.57	1.460, 1.666
BERT+CharSwap +WordNet		13.23	30.10	– , 1.757

Table 1: Sumerian-English Machine Translation. Here, S: Supervised, US: Unsupervised, SS: Semi-Supervised and HE: Human Evaluation. Each of the available values for the first three columns (BLEU) is compared with a value under HE (out of 3). *Number of shards used for FT.

	F1-Score
HMM	0.815
Rules + CRF	**0.991**
Bi-LSTM + CRF	0.763
FLAIR	0.499
RoBERTa	0.949

Table 2: POS Tagging for Sumerian. CRF with rules outperform large models like FLAIR and RoBERTa.

	F1-Score
HMM	0.656
Rules + CRF	0.913
Bi-LSTM + CRF	0.775
FLAIR	0.187
RoBERTa	**0.953**

Table 3: NER for Sumerian. RoBERTa performs best among others. Due to high character-level noise, FLAIR fails to generalise well.

GeForce RTX 2070 GPU, while the pre-training and fine-tuning of FLAIR, RoBERTa, and XLM on various data configurations were performed on 2 16 GB Nvidia V100 GPUs. We used development sets to tune the hyper-parameters for all our models, especially those for POS and NER. For RoBERTa and vanilla transformer, $N = 6$ encoder layers with $h = 16$ attention heads were used, while $N = 4$ and $h = 12$ was used for XLM. A beam-size of 5 was used for our FT experiments. Adam (Kingma and Ba, 2015) optimiser with a learning rate of 0.001, $\beta_1 = 0.90$, $\beta_2 = 0.98$ and a decay factor of 0.5 was used. Additional regularisation was done via Dropout and Attention Dropout (wherever applicable) layers with $p_{drop} = 0.1$. We used a batch size of 32 or 64 and an early stopping criteria based on the validation loss.

6 Results and Analysis

Sequence Labeling Tables 2 and 3 represent the metric scores of our different models for POS and NER tasks, respectively. CRF with domain-specific rules gives the best F1-score for the POS tagging task, even better than the complex RoBERTa and FLAIR language models which are the current state-of-the-art techniques for most languages. The prevalence of distorted words and short phrases in the corpora makes context learning difficult, although the domain-specific rules help learn short-term dependencies by learning feature weights.

RoBERTa performs well for both of the tasks, while being the best among others for NER (95.37 F1 score). To make the most out of the limited vocabulary and noisy text, we used Byte-Level BPE (Radford et al., 2019) to train the language model and further fine-tuned it on our POS and NER dataset with a batch size of 128. We also tried FLAIR language model across various word embeddings (character, Word2vec, FastText, GloVe) along with an additional CRF layer for both of the tasks. Although a high precision is observed using this approach, the F1 scores is seen to be significantly low due to low recall. In addition to the F1 metric we also

conducted human evaluation by language expert for the best performing models, out of randomly selected 76 (496 words) phrases, only 8 and 6 words were misclassified by NER and POS models, giving an error of 1.20 and 1.61%, respectively.

Machine Translation Table 1 summarises our results for all supervised, semi-supervised, and unsupervised techniques. Forward translation on vanilla transformer outperforms all other techniques by at least 2 BLEU. The variation of its performance with more monolingual source text is shown. The superior performance of `AllComp` over the other configurations in vanilla transformer signifies the value of both context and out-of-domain data together. Even though the XLM-based models show lower performance, it could be attributed to the lesser number of encoder layers and attention heads used for them. What is interesting to note, though, is the variation of its performance across various training strategies. We experiment across MLM and TLM (+ MLM) initialization for XLM, where the latter comfortably outperforms the former. We do not test with random initialization and CLM, following up from the conclusions made for NMT in Conneau and Lample (2019). Pre-training the XLM on augmented target-side text works surprisingly well. We note that using pre-training on `BERT` and `WordNet` augmentations results in better Unsupervised performance while introducing `CharSwap` improves the semi-supervised models. The human evaluation presented in the table was made by three Assyriologists, who rated 100 output examples for each model, on a scale of 3. A pairwise inter-annotator agreement of 0.673 (Cohen's Kappa) was observed.[6]

7 Interpretability Analysis

Oftentimes in case of Deep Learning Architectures, metric scores like Accuracy, F1 and BLEU are unable to portray the true behavior of the models. For languages like Sumerian, the human-understanding itself is scarce. Visualizing the representations and correlations made by the model could provide insights into which elements of the context can give additional information to support semantic analysis of the terms. Thus, we herein introduce a generalisable interpretability toolkit, *InterpretLR*, to interpret algorithms for **L**ow-**R**esource NLP and further apply it for the aforementioned tasks and models.

InterpretLR is primarily aimed at fabricating attribution saliency maps, i.e., tracing back the model output so as to assign an importance score to each input token, based on its 'influence' on that output. We do this using two kinds of interpretability techniques– gradient-based (Sundararajan et al., 2017; Simonyan et al., 2014; Shrikumar et al., 2017), and perturbation-based (Zeiler and Fergus, 2014; Castro et al., 2009).

Due to the inherently discrete nature of natural language text, the starting point for all our approaches is the embedding of the input sentence across the model to interpret. Most of our analysis is done for the encoder of the network architecture, thus analyzing the effect of different pre-training and fine-tuning techniques on how the model eventually represents the language attributes. We use the word 'Attribution' as a better-defined substitute for the 'Influence' measure of an input span of text on the output.
A part of our visual analysis is shown and elaborated here, while a complete analysis with all our models and layer-wise heat-maps is presented in the Appendix.

In Table 4a, we apply *InterpretLR* on 3 different configurations of XLM for a randomly chosen sentence from NMT's evaluation set. A human expert was asked to annotate the source sentence in accordance with the expected reference for each output token in the actual English translation, as shown in the first column. The highlighted visualizations for each of the 3 models were obtained using Integrated Gradients (Sundararajan et al., 2017) across the three input embeddings- token, position, and language. A lot of interesting observations could be made from these attributions.
Firstly, the named entity in the sentence *ur-{d}asznan* (*UrAnan*) has been wrongly translated by all the three models. Although this behavior is expected (learning the context of a named entity is extremely difficult without excessive supervision around the same, which is largely absent our training text) the models even largely fail to attend to the right words in the input.
Secondly, words like *rations*, *weavers* and *seal* which appear frequently in the parallel Ur III Admin corpora and have a contextual meaning attached to them, are translated perfectly by the models, this property is observed among these models

[6] Elaborate evaluation criteria mentioned in the Appendix.

Actual	**Human Expert**	Model-1	**Semi-Supervised DataAug XLM**	Model-2	**Unsupervised DataAug XLM**	Model-3	**Unsupervised `Orig` TLM XLM**
Output Word	Annotations	Output Word	Visualisations	Output Word	Visualisations	Output Word	Visualisations
barley	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	barley	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	Monthy	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	Basketoftablets	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e
rations	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	rations	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	rations	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	rations	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e
weavers	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	weavers	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	weavers	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	weavers	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e
under	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	under	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	from	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	255	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e
seal	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	seal	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	seal	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	seal	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e
of	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	of	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	of	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	of	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e
UrAnan	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	Lugalniglagare	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	Ninlil	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	weavers	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e
foreman	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	foreman	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	foreman	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e	female	#s sze-ba geme2 usz-bar kiszib3 ur-{d}asznan ugula #e

(a) MT- Selected output tokens for Sumerian Input text of "sze-ba geme2 usz-bar kiszib3 ur-dasznan ugula", which translates to "barley rations of the female weavers under seal of UrAnan the foreman".[7]

Actual	**Human Expert**	Model	**RoBERTa**
N	5(disz) gin2 ku3-babbar	N	5 (disz) gin 2 ku- 3 - babbar

(b) POS- With tagged word "ku3-babbar"

Actual	**Human Expert**	Model	**RoBERTa**
GN	mu ur-bi2-lum{ki} ba-hul	GN	mu ur - bi 2- lum { ki } ba - hul

(c) NER- With tagged word "ur-bi2-lumki"

Table 4: Highlighted attributions for randomly selected examples. **Green** and **Red** represent correct and wrong predictions, respectively, while Green and Red highlights represent positive and negative attributions, respectively.

in general. Even the unsupervised models that do not have access to the one-to-one mapping of the translation during training manage to infer these words from the appropriate context. It can be assumed that they learn the right representations of such tokens. But at the same time, there are instances like *sze-ba* (*barley*), which the two unsupervised models rightly refer to but do not give the right translations, which thus is a direct result of the absence of supervision.

Lastly, English words like *under*, *of* and *from* do not have any direct translations in Sumerian and are mostly inferred from the context, even by the human annotators. At such places, again, supervision might play a critical role as in the 4^{th} row of Table 4a. There are also instances like the 6^{th} row where the supervised model fails to attend to the right words, and the correct output word could very well be out of memorisation.

Tables 4b and 4c represent visualizations for two randomly selected phrases for our sequence labeling tasks, indicating the attributions for each sub-word for tagging the corresponding target word with their predicted labels. It can be observed from Table 4b that word *gin* (*unit*) and sub-word *ku*, are contributing to the attribution score positively, depicting positive model attribution to tag *ku3-babbar* as a Noun (N), whereas in Table 4c the sub-words *ur*, *hul* and *ki* are contributing *ur-bi2-lum*$\{ki\}$ to be tagged as the label GN (Geographical Name). As observed from the corresponding human annotation, *ur* and *ki* are the most associated for Geographical names and GNs are mostly followed by a verb part, which is *hul* (*destroy*) in this case. It can thus be inferred that RoBERTa identifies this correspondence well and makes the decision accordingly.

[7]The left-out tokens were rightly predicted by all the three models, with almost the same attributions.

8 Conclusion

In this work, we introduced the first information extraction and translation pipeline for Sumerian cuneiform. We first undertook the tasks of POS Tagging and NER, where we observed that *deeper is not necessarily better*. A simple CRF model with well-defined rules outperformed the large language model RoBERTa for POS Tagging. Further, for machine translation we overcame unprecedented challenges pertaining to lack of in-domain text, sparse sentence formation, and incoherence. We found that using out-of-domain text along with specific data-augmentation can have huge impacts in a low-resource setting. All components of this work are generalisable to other low-resource languages, including *InterpretLR*, and we open way to future research in this direction.

Acknowledgments

The authors would like to acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility in carrying out this work (http://dx.doi.org/10.5281/zenodo.22558). We would like to thank our collaborators at the Cuneiform Digital Library Initiative (CDLI; https://cdli.ucla.edu) and from the Machine Translation and Automated Analysis of Cuneiform Languages (MTAAC; https://cdli-gh.github.io/mtaac/). We would also like to thank CDLI and the Electronic Text Corpus of Sumerian Literature (ETCSL) for providing the data for our experiments. This work was partly undertaken during the Google Summer of Code (GSoC) program, 2020. CDLI has been supported by GSoC, where aspects of machine translation have been addressed by several students since 2018. We are thankful to Ilya Khait and Bertrand Lafont for their assistance with the human evaluations for machine translation and sequence labeling. We are grateful to Orhan Firat for insightful discussions and multiple rounds of reviews during the pre-submission mentoring phase of ACL SRW that greatly shaped this manuscript.

References

Alan Akbik, Tanja Bergmann, Duncan Blythe, Kashif Rasul, Stefan Schweter, and Roland Vollgraf. 2019a. FLAIR: An easy-to-use framework for state-of-the-art NLP. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations)*, pages 54–59, Minneapolis, Minnesota. Association for Computational Linguistics.

Alan Akbik, Tanja Bergmann, and Roland Vollgraf. 2019b. Pooled contextualized embeddings for named entity recognition. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, pages 724–728. Association for Computational Linguistics.

Alan Akbik, Duncan Blythe, and Roland Vollgraf. 2018. Contextual string embeddings for sequence labeling. In *Proceedings of the 27th International Conference on Computational Linguistics, COLING 2018, Santa Fe, New Mexico, USA, August 20-26, 2018*, pages 1638–1649. Association for Computational Linguistics.

Franck Burlot and François Yvon. 2018. Using monolingual data in neural machine translation: a systematic study. In *Proceedings of the Third Conference on Machine Translation: Research Papers*, pages 144–155, Brussels, Belgium. Association for Computational Linguistics.

Javier Castro, Daniel Gómez, and Juan Tejada. 2009. Polynomial calculation of the shapley value based on sampling. *Comput. Oper. Res.*, 36(5):1726–1730.

Alexis Conneau and Guillaume Lample. 2019. Cross-lingual language model pretraining. In *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, pages 7057–7067.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, pages 4171–4186. Association for Computational Linguistics.

Sergey Edunov, Myle Ott, Michael Auli, and David Grangier. 2018. Understanding back-translation at scale. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, October 31 - November 4, 2018*, pages 489–500. Association for Computational Linguistics.

Robert K. Englund. 2009. The smell of the cage. `https://cdli.ucla.edu/pubs/cdlj/2009/cdlj2009_004.html`. Online; accessed 2009.

Di He, Yingce Xia, Tao Qin, Liwei Wang, Nenghai Yu, Tie-Yan Liu, and Wei-Ying Ma. 2016. Dual learning for machine translation. In *Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5-10, 2016, Barcelona, Spain*, pages 820–828.

Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Comput.*, 9(8):1735–1780.

Ari Holtzman, Jan Buys, Maxwell Forbes, Antoine Bosselut, David Golub, and Yejin Choi. 2018. Learning to write with cooperative discriminators. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 1: Long Papers*, pages 1638–1649. Association for Computational Linguistics.

Zhiheng Huang, Wei Xu, and Kai Yu. 2015. Bidirectional lstm-crf models for sequence tagging. *arXiv preprint arXiv:1508.01991*.

Diederik P. Kingma and Jimmy Ba. 2015. Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.

Guillaume Klein, Yoon Kim, Yuntian Deng, Jean Senellart, and Alexander Rush. 2017. OpenNMT: Open-source toolkit for neural machine translation. In *Proceedings of ACL 2017, System Demonstrations*, pages 67–72, Vancouver, Canada. Association for Computational Linguistics.

John D. Lafferty, Andrew McCallum, and Fernando C. N. Pereira. 2001. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the Eighteenth International Conference on Machine Learning (ICML 2001), Williams College, Williamstown, MA, USA, June 28 - July 1, 2001*, pages 282–289. Morgan Kaufmann.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized BERT pretraining approach. *CoRR*, abs/1907.11692.

Yudong Liu, Clinton Burkhart, James Hearne, and Liang Luo. 2015. Enhancing sumerian lemmatization by unsupervised named-entity recognition. In *NAACL HLT 2015, The 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Denver, Colorado, USA, May 31 - June 5, 2015*, pages 1446–1451. The Association for Computational Linguistics.

Liang Luo, Yudong Liu, James Hearne, and Clinton Burkhart. 2015. Unsupervised sumerian personal name recognition. In *Proceedings of the Twenty-Eighth International Florida Artificial Intelligence Research Society Conference, FLAIRS 2015, Hollywood, Florida, USA, May 18-20, 2015*, pages 193–198. AAAI Press.

George A Miller, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine J Miller. 1990. Introduction to wordNet: An on-line lexical database. *International journal of lexicography*, 3(4):235–244.

Myle Ott, Sergey Edunov, Alexei Baevski, Angela Fan, Sam Gross, Nathan Ng, David Grangier, and Michael Auli. 2019. fairseq: A fast, extensible toolkit for sequence modeling. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Demonstrations*, pages 48–53. Association for Computational Linguistics.

Émilie Pagé-Perron, Maria Sukhareva, Ilya Khait, and Christian Chiarcos. 2017. Machine translation and automated analysis of the Sumerian language. In *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, pages 10–16, Vancouver, Canada. Association for Computational Linguistics.

Ravneet Punia, Niko Schenk, Christian Chiarcos, and Émilie Pagé-Perron. 2020. Towards the first machine translation system for sumerian transliterations. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 3454–3460.

Alec Radford, Jeff Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. 2019. Language models are unsupervised multitask learners.

Mike Schuster and Kuldip K. Paliwal. 1997. Bidirectional recurrent neural networks. *IEEE Trans. Signal Process.*, 45(11):2673–2681.

Rico Sennrich, Barry Haddow, and Alexandra Birch. 2016. Improving neural machine translation models with monolingual data. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, ACL 2016, August 7-12, 2016, Berlin, Germany, Volume 1: Long Papers*. The Association for Computer Linguistics.

Avanti Shrikumar, Peyton Greenside, and Anshul Kundaje. 2017. Learning important features through propagating activation differences. In *Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, NSW, Australia, 6-11 August 2017*, volume 70 of *Proceedings of Machine Learning Research*, pages 3145–3153. PMLR.

Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman. 2014. Deep inside convolutional networks: Visualising image classification models and saliency maps. In *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Workshop Track Proceedings*.

Kaitao Song, Xu Tan, Tao Qin, Jianfeng Lu, and Tie-Yan Liu. 2019. MASS: masked sequence to sequence pre-training for language generation. In *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, volume 97 of *Proceedings of Machine Learning Research*, pages 5926–5936. PMLR.

Mukund Sundararajan, Ankur Taly, and Qiqi Yan. 2017. Axiomatic attribution for deep networks. In *Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, NSW, Australia, 6-11 August 2017*, volume 70 of *Proceedings of Machine Learning Research*, pages 3319–3328. PMLR.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all

you need. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, pages 5998–6008.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, and Jamie Brew. 2019. Huggingface's transformers: State-of-the-art natural language processing. *CoRR*, abs/1910.03771.

Lijun Wu, Yiren Wang, Yingce Xia, Tao Qin, Jianhuang Lai, and Tie-Yan Liu. 2019. Exploiting monolingual data at scale for neural machine translation. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 4207–4216, Hong Kong, China. Association for Computational Linguistics.

Matthew D. Zeiler and Rob Fergus. 2014. Visualizing and understanding convolutional networks. In *Computer Vision - ECCV 2014 - 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I*, volume 8689 of *Lecture Notes in Computer Science*, pages 818–833. Springer.

Jiajun Zhang and Chengqing Zong. 2016. Exploiting source-side monolingual data in neural machine translation. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1535–1545, Austin, Texas. Association for Computational Linguistics.

Barret Zoph, Deniz Yuret, Jonathan May, and Kevin Knight. 2016. Transfer learning for low-resource neural machine translation. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, EMNLP 2016, Austin, Texas, USA, November 1-4, 2016*, pages 1568–1575. The Association for Computational Linguistics.

A Detailed Evaluation and Analysis

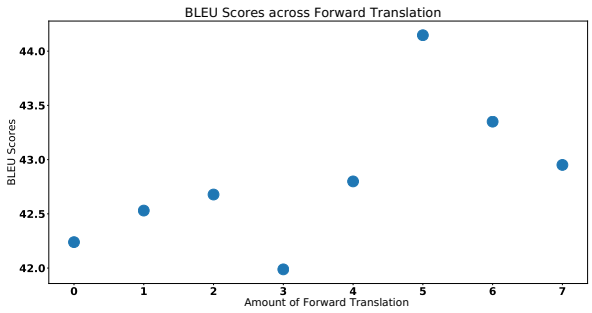
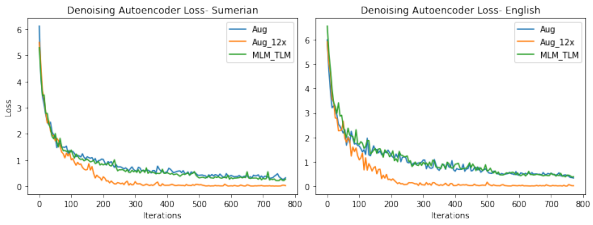


Figure 5

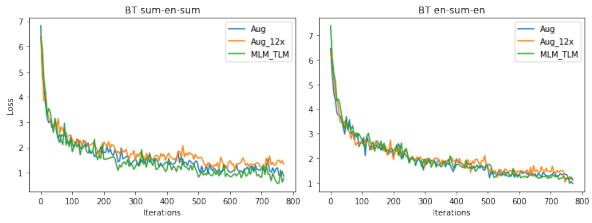
Forward Translation with Vanilla Transformer gave the best results for Sumerian-English Neural Machine Translation. Figure 5 shows the variation of the BLEU score with the amount of source monolingual data used. Here, the X-Axis represents the number of shards used, with each shard consisting of 80K sentences. It can be observed that the translation accuracy is not linear with the amount of text used.

Figure 6 shows the variation of several performance metrics during the Unsupervised fine-tuning of various XLM configurations. The comparison is made between XLM pre-training without any data augmentation (MLM_TLM), with one augmentation (Aug) and with all three augmentations (Aug_12x). It can be seen from Figure 6a that an XLM pre-trained on the Aug_12x configuration converges the fastest among the others, in terms of the main Denoising Auto-encoding Loss. It can also be observed that the curve corresponding to this configuration is much smoother than the others, which shows a positive regularizing effect of a better weight initialisation (through appropriate pre-training). A similar pattern is observed for the validation accuracy across the epochs as shown in Figure 6c, although, the trend of Back Translation loss remains mostly inseparable for the three configurations.

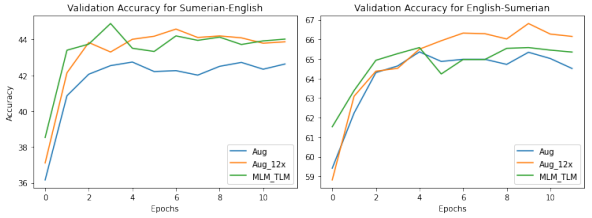
Table 5 depicts the net percentage error found by an human expert on the POS and NER results for the entire evaluation set across the best performing model. Table 6 and 7 represents the detailed results of POS and NER models. It can be observed from the tables, that although CRF and RoBERTa models gave the best results, FLAIR language model along with character embeddings also gave high precision for both of the tasks.



(a) Denoising Auto-encoder Loss (AE Loss) variation across the 1st Epoch



(b) Back Translation Loss variation in XLM across the 1st Epoch



(c) MT accuracy across a number of training epochs

Figure 6: Quantitative comparison of various models during Unsupervised MT fine-tuning

	POS error (in %)	NER error (in %)
Human Evaluation	1.61	1.20

Table 5: Human Evaluation for POS and NER

B Extended Interpretations

Here we present the interpretability analysis across a larger set of models and visualisations. We use and compare the different algorithms across layer-level, gradient-based, and perturbation-based techniques to obtain the attributions.

Figure 7 visualises the Multi-head Self Attention (MHSA) using Layer Conductance Dhamdhere, Sundararajan, and Yan 2018) across the 4 encoder layers we employ in XLMs[8]. The first two output tokens *barley* and *female* are known to be one-on-one mapping between the input words of *sze-ba* and *geme2* respectively. While the third output token *barley* is not a direct translation and

[8]The supervised version of the augmented pre-training is used here.

	Part of Speech Tagging		
	Precision	Recall	F1-Score
HMM	0.857	0.794	0.815
Rules + CRF	**0.994**	**0.989**	**0.991**
BBi-LSTM + CRF	0.852	0.710	0.7631
FLAIR	0.9323	0.4766	0.4999
RoBERTa	0.9500	0.9489	0.9495

Table 6: POS Tagging Models for Ur III Sumerian Text

	Named Entity Recognition		
	Precision	Recall	F1-Score
HMM	0.810	0.599	0.656
Rules + CRF	0.916	0.910	0.913
Bi-LSTM + CRF	0.864	0.704	0.775
FLAIR	**0.9562**	0.1817	0.1873
RoBERTa	0.9540	**0.9534**	**0.9537**

Table 7: NER Models for Ur III Sumerian Text

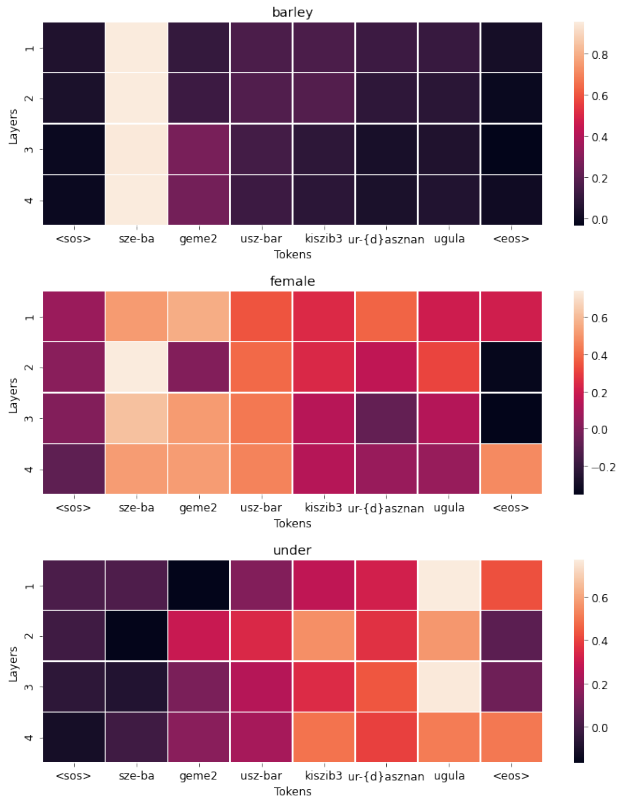


Figure 7: Layer Conductance across MHSA Layers

is needed to be inferred from context.

Figure 9a represents the attribution heat-map when gradient-normalisation saliency (Simonyan, Vedaldi, and Zisserman 2013) is used. Being one of the most conventional techniques for finding attribution, it is more prone to inconsistent interpretations. Whereas, the attribution heat-map in Figure 9b represents the Integrated Gradients (IG) (Sundararajan, Taly, and Yan512017) approach. Being a path-based technique, which measures the gradient attribution relation using a straight-line path from a baseline (usually all-zeros), to the given input, it is much more robust and stable.

Even though the gradient-based methods are much faster than perturbation-based methods, we observe that the heavy dependency of IG on hyper-parameters like the number of input steps to be considered when going from a baseline to the actual input, n_steps, to be a major setback. The final attribution is generally found out after integrating (or summing) over the attributions of these sub-steps. We found that the attributions do not change when going beyond $n_steps = 250$, thus, we experiment by varying it between 10 to 250. We observe that there is no ideal value of n_steps, IG's faithfulness to the model varies largely over this range. For some inputs, the best value is $n_steps = 50$ while for others $n_steps = 250$ is the most ideal. We judge this by considering how much the attribution is given to *sos* and *eos* tokens for each output token. Thus, based on both *plausibility* and *faithfulness*. We use $n_steps = 50$ for obtaining the heat-maps in Figure 9b.

Figure 10 represents the visualization for our sequence labeling tasks. It indicates two major things, 1) the effect of words, sub-words (depends on tokenization) on tagging the target word and 2) the effect of 6 transformer encoder layers. We created the hook on embeddings of RoBERTa with layer IG and obtained the visualizations for how each sub-word is contributing to tag the target word. Similarly, to obtain the heat-map we created the hook on RoBERTa embeddings and used the Layer Conductance.

From Figure 10a it can be observed that *ku* and *du* contribute the most to the attribution scores for tagging *ku3-babbar* and *ba-du3* as a Noun (N) and Verb (V), respectively. From the heat-maps it is also noted that *ku* shows the effect on all 6 layers whereas in second example effects are majorly due to the initial transformer layers. Similarly in the Figure 10b *ur* and *lugal* are the most effective sub-words to tag *ur-bi2-lumki* and *lugal-tesz2-mu* as

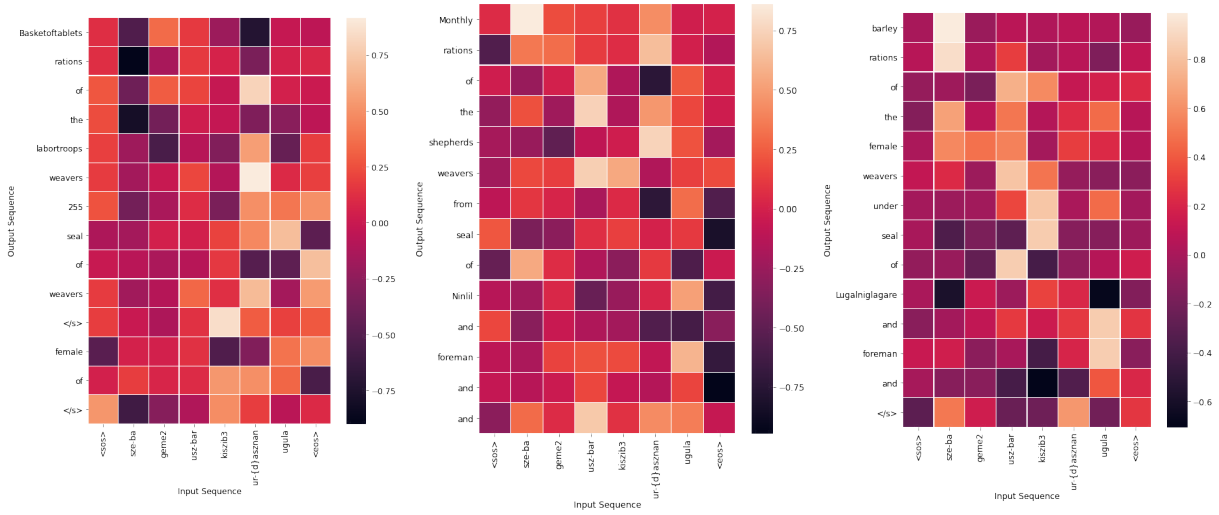


Figure 8: Feature Ablation in *InterpretLR*

GN (Geographical Name) and PN (Personal Name) respectively. It is also interesting to note that both of these sub-words have a very positive impact in the initial layers but are contributing oppositely in the last layer.

B.1 Human Evaluation

The scoring by human experts was done independently for each result according to the following criterion:

• **3 (good)**: interpretable in the correct meaning by a native speaker of English; (almost) no incorrectly translated content word (e.g., tolerant against some errors in word order, but not in incorrect words).

• **2 (helpful)**: partially distorted, but interpretable with some context information (tolerant against errors in word order and against incorrect function words).

• **1 (incorrect)**: contains incorrectly translated content words and/or is un-interpretable.

C Rules for POS Tagging and NER

We used certain language-specific rules to assist CRF for the sequence labeling tasks. The rules were identified by human experts and some of them are as mentioned here:

• A word starting with “ur-”, “lu2-”, or “dumu” is most likely to be a personal name.

• If a word is followed by “mu”, then the next phrase denotes a year name.

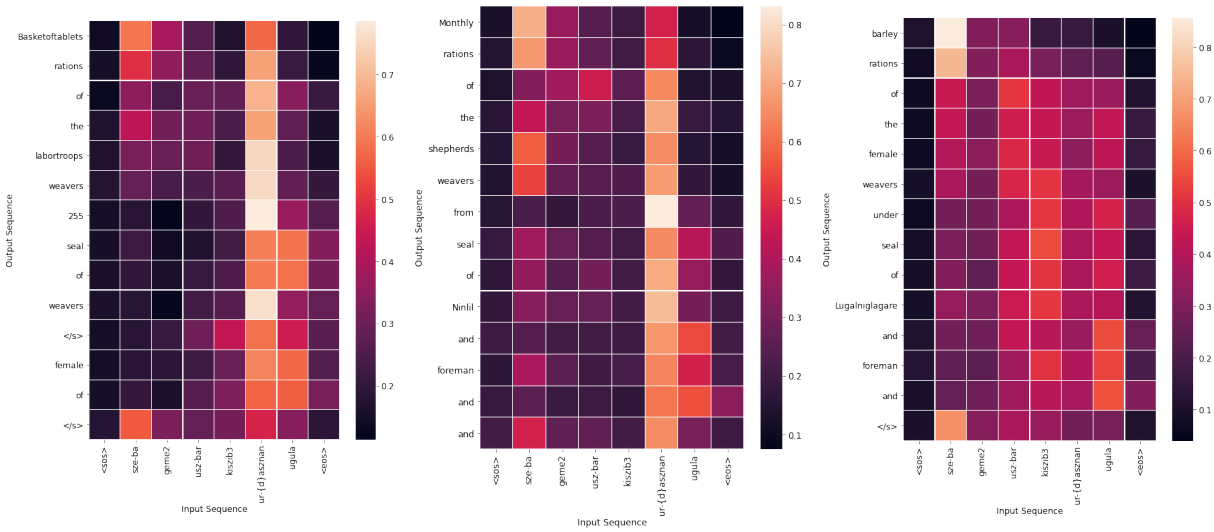
• If a word is followed by “iti”, it denotes a month name.

• Words containing “ki” are mostly associated with geographical names (GN).

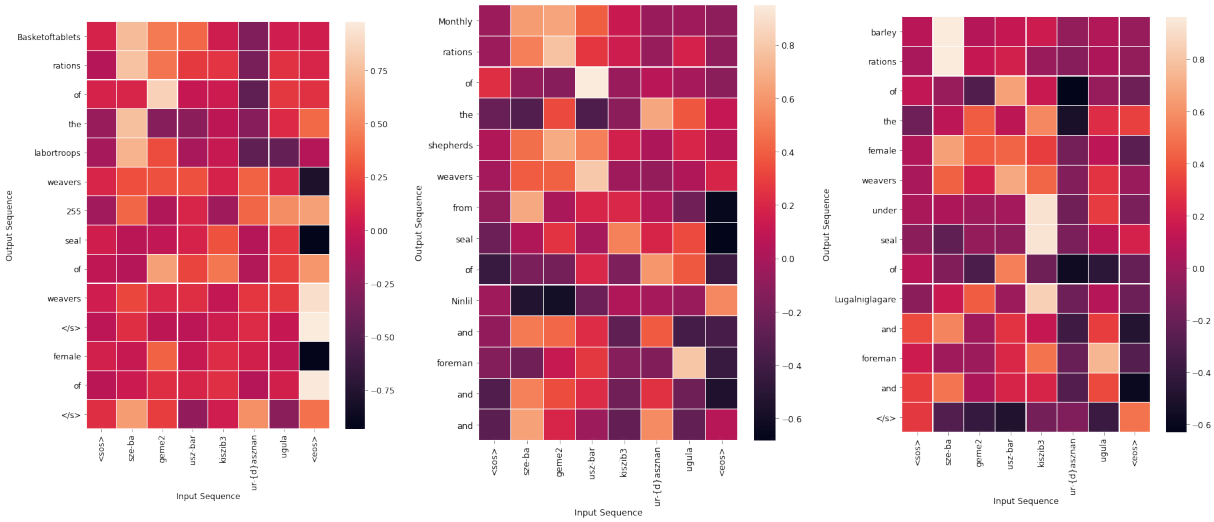
• Words ending with part “-hul” majorly denotes verbs.

• Words containing “{d}” denotes either personal name (PN) or divine name (DN).

• A word followed by “gin” (*unit*) majorly replicate a noun.



(a) Grad-Norm Saliency



(b) Integrated Gradients

Figure 9: Comparing different gradient-based approaches used in *InterpretLR*



(a) POS Tagging (b) NER

Figure 10: *InterpretLR* on RoBERTa for Sequence Labeling

Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Hydrogen Production from Biogas: Methods and Economic Analysis

Ravindra Kumar
Delhi Technological University
Anil Kumar (anilkumar76@dtu.ac.in)
Delhi Technological University https://orcid.org/0000-0002-2846-4347
Amit Pal
Delhi Technological University

Research Article

Keywords: Hydrogen production, Biogas, Anaerobic, digestion, Reforming method

DOI: https://doi.org/10.21203/rs.3.rs-457836/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

In present days, H_2 is mainly produced via natural gas from thermochemical methods. The importance of renewable H_2 production has increased due to energy issues, global environmental, fossil fuel depletion, and pure hydrogen demand. Biogas is produced from anaerobic digestion or carbon-based waste fermentation methods. It might be a substitute resource to the creation of hydrogen. Methane (CH_4) and carbon dioxide (CO_2) are two significant constituents of biogas, and these are multipurpose feedstock materials for the generation of valuable chemicals and fuels. This paper presents major hydrogen generation techniques from biogas reforming in detail and a short discussion of biogas reforming's economic analysis. The effective operating parameters such as reaction temperature, gas hourly space velocity and O_2/CH_4 ratio are also discussed. Biogas has more potential; however, it needs more research on some issues such as deactivation of the catalyst, purification of biogas and removable of H_2S. Various reforming methods can be nominated for H_2 production depend on biogas composition, the need for hydrogen purity, investment availability and quantity of the desired hydrogen.

1. Introduction

Growing concern about worldwide climate issues and diminution of fossil fuels, the utilization of renewable energy resources has been more focussed (Avraam et al. 2010). A large scale of H_2 production developed via reforming techniques, mainly natural gas (CH_4) and light hydrocarbons (HCs), used in the chemical industry field (Alves et al. 2013). Hydrogen is abundant and is the most available in renewable energy. Furthermore, only water vapor is produced from the combustion of hydrogen. At present, biogas is an important group from renewable energy sources.

Biogas can be an important path of hydrogen production. It is a combination of dissimilar gases developed via the anaerobic breakdown or fermentation of carbon-based waste material in the existence of microbes (Verma and Samanta 2016). The product of biogas, mainly carbon dioxide (CO_2) and methane (CH_4), associated with touches of additional gases like hydrogen (H_2), ammonia (NH_3), hydrogen sulfide (H_2S), vapor water (H_2O), nitrogen (N_2), and oxygen (O_2) (Alves et al. 2013). The unwanted component of biogas is hydrogen sulfide (H_2S). It is corrosive gas and sustains to damage accessories and equipment during the process of producing energy. Another impurity, like CO_2 and humidity, also decreases biogas' calorific value through explosion for thermal plants (Horikawa et al. 2004). Biogas generally carries different gases as a basis of volume percentage by CH_4 (55–65), CO_2 (35–45) and N_2 (1) in sewage digesters while CH_4 (60–70), CO_2 (30–40) and N_2 (1) in biological waste digesters and CH_4 (45–55), CO_2 (30–40) and N_2 (5–15) in landfills (Lau et al. 2011). It is usually generated in biowaste digesters, sewage sludge and landfills that contain agricultural wastes, crop residue, animal wastes and animal dung etc. (Chattanathan et al. 2014). Some important chemical composites involve in biogas as basis of volume percentage listed in Table 1.

Table 1
Chemically structure of biogas (Alves et al. 2013; Verma and Samanta 2016; Nalbant and Colpan 2020)

Composite	Percentage
CH_4	55–70 (volume %)
CO_2	30–45 (volume %)
H_2S	500–4000 (ppm)
NH_3	100–800 (ppm)
H_2	Below 1 (volume %)
N_2	Below 1 (volume %)
H_2O	Below 1 (volume %)
O_2	Below 1 (volume %)

H_2 is the simplest element, an atom containing only one proton and one electron. It is a colorless and odorless gas. It is also the most plentiful component in the universe. It is combined with additional chemical substances and detected as part of a different element such as alcohol, hydrocarbon, and water. It is also available in inherent biomass, which contains animals and plants (Kalamaras and Efstathiou et al. 2013). The main benefits of hydrogen are net zero emission of greenhouse gases and high energy content compared to other known substances (LHV;120 MJ/kg and HHV;142 MJ/kg) (Ouyang et al. 2013).

In the present scenario, hydrogen is produced from fossil fuels around 96% and 1% from biomass (Sinha and Pandey 2011). Currently, application H_2 as raw material is widely used in a combustion engine, fuel cell, fertilizer industries, chemical industry, pharmaceutical industry, hydrogenation process, Fischere-Tropsch synthesis, production of ammonia and ammonia, among others (Armor 1999).

Overall, biogas reforming is a favorable and attractive method to generate clean hydrogen gas. The main biogas benefits for reforming purpose in H_2 production are decreased greenhouse gas emissions, local availability, reliability and better environment, and economic (Nalbant and Colpan 2020). The key motive of this paper is to express the numerous biogas reforming technique for H_2 production and its economic analysis.

2. Reforming Of Biogas For H2 Production

There are several approaches for the invention of H_2. It could be divided into two classes as traditional and non-traditional ways. In traditional methods, hydrogen is obtained from hydrocarbon reforming as

natural gas (CH_4) and hydrocarbon pyrolysis. In non-traditional methods or alternative methods that can use renewable sources (Nikolaiis and Poullikkas 2017).

Interestingly, methane is part of the biogas component, and biogas is generated from waste materials as adopted the renewable source. Therefore, biogas reforming seems to renewable energy from conventional techniques for the production of hydrogen (Nikolaiis and Poullikkas 2017).

To produce pure hydrogen, additional steps are included in any reforming process. A conversion reactor is utilized for converting the CO into CO_2 by using shift reaction (Alves et al. 2013). The final process for hydrogen purification is done in a separate unit by pressure swing adsorption (PSA) (Nikolaiis and Poullikkas 2017). H_2 could be obtained through the methane or reforming of biogas in various temperature limits (873–1273K) over reversible reactions and exothermic or endothermic. The reforming reactions are done by low pressure and high temperature. In several kinds of reforming methods, biogas responds along with representatives like oxygen or air and steam to generate syngas with H2 and other gases (Galvagno et al. 2013).

There are some purification techniques available for removing toxic and undesirable substances present in biogas: physiochemical and biological treatment (Alves et al. 2013). Figure 1 presents the important steps for the production of H_2 from biogas.

Consequently, the biogas reforming reactions are parallel to natural gas (CH_4) reforming. Table 2 illustrates the chemical reactions formed in reforming methods and its enthalpy of reaction at 298K and 1 atm. There are major reforming techniques for H_2 production from biogas or methane as follows.

Table 2 Chemically reaction involved in reforming methods (Alves et al. 2013; Verma and Samanta 2016)

Reaction Equation	Name of reaction	Type of reaction	Thermal reaction, ΔH_{298} K (kJ/mol)
Eq. 1	SR	$CH_4 + H_2O \rightarrow CO + 3H_2$	206.20
Eq. 2	Gas-shift reaction	$CO + H_2O \rightarrow CO_2 + H_2$	−41.20
Eq. 3	Combined reaction	$CH_4 + 2H_2O \rightarrow CO_2 + 4H_2$	165
Eq. 4	Methane cracking	$CH_4 \rightarrow C + 2H_2$	74.9
Eq. 5	Boudouard reaction	$2CO \rightarrow C + CO_2$	−172.40
Eq. 6	Reduction of CO	$CO + H_2 \rightarrow C + H_2O$	−131.30
Eq. 7	POR	$CH_4 + 0.5O_2 \rightarrow CO + 2H_2$	−35.60
Eq. 8	Complete oxidation	$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$	−801.70
Eq. 9	ATR	$CH_4 + 0.5xO_2 + yCO_2 + (1-x-y)H_2O \rightarrow (y+1)CO + (3-x-y)H_2$	≈ 0
Eq. 10	DR	$CH_4 + CO_2 \rightarrow 2CO + 2H_2$	247.40
Eq. 11	DOR	$CH_4 + \beta CO_2 + (1-\beta)/2\ O_2 \rightarrow (1+\beta)CO + 2H_2$	$(285\beta - 38)$ $0 \leq \beta \leq 1$

2.1 Steam Reforming (SR)

This process is done by methane and water (H_2O) vapor with the catalyst that generates syngas (mixture of CO and H_2, Eq. 1, Table 2). Methane is required a high temperature for reacting. The reaction of this process is extremely endothermic and needs a high temperature of nearly 650–850°C to attain a high yield of H_2 in between 60 to 70% (Verma and Samanta 2016). The ratio H_2/CO is three that means the higher yield of H_2 nearly 75%. High temperature can also be reduced by using a suitable catalyst (Shanmugam et al. 2018).

This technique consists of a steam reformer, water gas shift reactor, and direct steam reforming reactions inside the reactor. Initially, methane responds with water in the steam reformer, and the syngas and CO and H2 are formed in the existence of a suitable catalyst. Then, the syngas is fed in the WGS reactor after cooling in a range of 300–500°C. CO reacts through the water to decrease the CO content of syngas and developed the H_2 and CO_2 (Eq. 2). High conversion of CO content and low kinetic reactions were done by high-temperature WGS reactor. While low CO conversion and high kinetic reactions are done by minimum temperature water-gas shift reactors (Minh et al. 2018). Fe, Cu, Fe-Pd and Mo alloys are mostly used catalysts in water gas shift (Verma and Samanta 2016).

In addition, combined reaction (Eq. (3)), also known as direct steam reforming reaction, is the addition of Eqs. 1 and 2. To yield pure hydrogen, need a separate unit for example, membrane or PSA (pressure swing adsorption), remove all gases as CO_2, CO, and other gases except H_2. Figure 2 shows the reactant and products for the steam reforming process.

The working conditions of steam reforming lead to corresponding reactions that reason carbon formation occur over the surface of catalytic as methane cracking reaction (Eq. (4)), Boudouard reaction (Eq. (5)), and reduction of CO reactions (Eq. (6)). When carbon development seems like the form of nanotubes, it offers a lower distribution of carbon at the surface of the catalyst. Consequently, the catalyst activity is conserved for a long duration (Gregorie et al. 2020).

The possibility of carbon formations is a serious issue since coke deposits and leading deactivation on catalyst surfaces such as Ni-catalyst surface and solid material of carbon; denoted by C. The basic arrangements can control the deposition of coke as if they promoted carbon gasification in water improving and its properties of adsorption (Nawfal 2015). The most popular catalyst is Ni-based on ceramic supports for this process due to its affordable, easily available, and high-quality catalytic properties. There are several catalysts established as Ni-based to improve system performance in the term of selectivity and reactivity (Nawfal 2015).

The Combine components to develop catalysts with oxides that are basic properties such as potassium, calcium and magnesium have been studied. Different catalysts such as pd and Pt-based catalyst have mere stability about coke formation (Alves et al. 2013). Ruthenium-based catalyst is also mostly used in the methane steam reforming method due to its more activity, the high selectivity of H_2 and outstanding stability (Nawfal 2015). Another way to decrease the deposition of coke as a result of a reaction (Eq. 6) is to perform under the condition of high partial pressure of water to shift the equilibrium in the right

direction and control the carbon formation (Noh et al. 2019). Some other studies explain significant development in separation unit process through the selective membrane reactors or filters (Mahecha-Botero et al. 2009).

All chemical reactions are done in a single vessel from a membrane reactor. To membrane system, both the reactions such as steam methane reforming reaction and shift reaction (Eqs. 1 and 2) co-occur inside the reactor, which comprises a catalyst bed. Additionally, membrane offers the good ability to shift the reaction towards a chemical equilibrium that improves hydrogen production and performs the temperature nearly below 500°C for reforming reaction in the direction of chemical equilibrium (Lin et al. 2003; Sato et al.2010).

2.2 Partial Oxidation Reforming (POR)

Partial oxidation reforming technique (Eq. 7, Table 2) is another alternative path to produce H_2 from methane reforming. This technique reduces energy consumption/cost due to the exothermic nature and is contrary to the steam reforming process that endothermic nature (Verma and Samanta 2016).

A methane (CH4) molecule is incompletely oxidised to produce H_2 and CO (known as syngas). The ratio of H_2/CO is two that means yields of hydrogen around 67%. The process needs a temperature range (700-900^oC) (Patinvoh et al. 2017). The molar ratio of O_2/CH_4 is a key parameter to control this process since the ratio of O_2/CH_4 more than 0.50 leads to perfect combustion of CH_4 (Eq. 8, Table 2) (Lin et al. 2003).

Since CO selectivity slightly decreases, CH4 reacts with O2 and obtains CO_2 that permits complete combustion and sudden temperature rise. That reason initiated a hot spot in the reactor bed that deposits coke formation over the surface of the catalyst (Patinvoh et al. 2017). Researchers have been studied the activity and stability of numerous catalysts with the purpose of POR of biogas. Temperature range more than 800°C, catalyst as a solid solution of (Ca–Sr–Ti–Ni), (Ni–Mg–Cr–La–O), (NiO–MgO), and mixed metal oxides and they are used as higher resistance of coke development (Verma and Samanta 2016; Nalbant and Colpan 2020; Gregorie et al. 2020).

2.3 Auto-Thermal Reforming (ATR)

This process is the union approach of SMR and POR and endothermic in nature. This process reduces steam reforming challenges due to its endothermic process and difficulty producing low yield of hydrogen of partial oxidation reforming (Gregorie et al. 2020).

The summation of reactions is expressed by Eqs. 3,7 and 8 (Table 2). It also happens in CO_2, as denoted by Eq. 9 (Table 2) (Araki et al. 2009). When the steam and oxygen are provided into the reformer, reason reforming and oxidation reactions occur parallelly, become self-sustainable the whole process, and require less energy. (Nikolaiis and Poullikkas 2017; Verma and Samanta 2016). This process produces the maximum hydrogen yield of approximate 74% (about ratio of H_2/CO is 2.8) (Nikolaiis and Poullikkas 2017).

‘Autothermal’ means that no need to supply from an external source for the process. ATR process stabled the heat requirement of the endothermic process through the heat released by the exothermic process and adopted it as an effective system (Nahar et al. 2017). ATR has more advantages, such as fast reactor processing, higher ratio of H_2/O_2 compared with POR, temperature control of processing, and minimize the hot spots that prevent the deactivation of catalysts (Verma and Samanta 2016; Patinvoh et al. 2017).

To obtain the H_2 through the biogas, the various the ratios range for H_2O/CH_4, O_2/CH_4, and H_2/CO are taken as 1.0–2.5, 0.25–0.55, and 2.0–3.5, respectively (Verma and Samanta 2016; Patinvoh et al. 2017; Mosayebi et al. 2012; Cai et al.2006).

The maximum efficiencies of plant to biogas reforming with ATR method are about 75%, range of temperature 500 -700^oC and the ratio of O/C close to 0.8–0.9 (Rau et al. 2019).

2.4 Dry Reforming (DR)

In this approach, CH_4 reacts with CO_2 and obtains the mixture of CO and H_2 (Syngas) (Verma and Samanta 2016), refer Eq. 10 Table 2. This process requirements temperature more than 640^oC and work on the endothermic process so need the heat of external source. Biogas is very useful for this process because it has two basic compounds available as CH_4 and CO_2 (Gregorie et al. 2020). This reaction of this process can be attractive an environmental point of view because they are strong greenhouse gases (CH_4 and CO_2), but this process can consume it. Contrarily, the rise in CO_2 emissions due to the proven external heat (other burned fuels) to endothermic reaction (Nalbant and Colpan 2020). Dry reforming uses the ratio of H_2/CO nearly 1 and attaining the yield of H_2 about 50%. The H_2/CO ratio built this reaction attractive about requirements to further use synthesis gas as Fischer–Tropsch synthesis (Gregorie et al. 2020).

Based on the previous finding, the main reaction (Eq. 10 in Table 2) could modify by competing along with some parallel reactions, which shifts the CO_2 conversion in (CH_4) methane. These important reactions are mainly such as the reaction of reverse gas water shift (Eq. 2 reverse), Boudouard reaction (breakdown of CO_2, Eq. 5 in Table 2), and decomposition of methane (Eq. 4, Table 2). DR method attains the ratio of CH_4/CO_2 close to 1-1.5 (Verma and Samanta 2016; Serrano-Lotina et al. 2012). The essential issue of DR method is the coke establishment that causes the deactivation of catalyst and blockage of the pipes and reactors (Alves et al. 2013). Pt-, Ru- and Rh-based catalysts improve the H_2 generation rate, but they are more expensive, very limited availability, and not established at an industrial scale. Generally, common catalysts such as Co- and Ni- based catalysts are more used (Cai et al. 2006).

2.5 Dry Oxidation Reforming (DOR)

This method has been established for governing the congestion of carbon formation over catalyst surface. It is a joined approach of DR and POR process. This process decreases the energy demand of reaction because the exothermic for partial oxidation releases high temperature that is beneficial to the

endothermal and dry oxidation reforming process. The ratio of H_2/CO also increases and achieving yield of H_2 approximately 60% (Gregorie et al. 2020).

There are additional advantages for this process where parallel feeding of CH_4 with CO_2 and O_2 such as the improved conversion of methane, decreased total energy of the process, higher catalyst activity, improved H_2 yield at lower temperatures, and enhanced resistance of deactivation (Chen et al. 2010). Eq. 11 represents the DOR process where β is stoichiometric ratio of CO_2 fed with the traditional DR (Table 2). If β value is 1, the reaction would be non-oxidative type DR reaction; now, if β value is 0, the reaction would be partial oxidation hence, as the β value must be in between 0–1(Alves et al. 2013; Verma and Samanta 2016; Nalbant and Colpan 2020).

In the DOR process, the process of changing parameters like reaction temperature and O_2 concentration feed supports the control of the ratio of H_2/CO and the appearance of reaction such as exothermic or endothermic. At a specific temperature, the nature of exothermic is increased by the increasing concentration of O_2. As known, the reaction of POR (Eq. 7) is exothermic and reaction of DR (Eq. 10) is endothermic by nature. Therefore, DOR method can be used as a preferable substitute over the convention DR method using various molar ratios of oxygen. Thus, the production of syngas from biogas through this method has improved energy performance and requires less exterior energy (Avraam et al.2010; Alves et al. 2013).

3. Economic Analysis Of Biogas Reforming

H_2 generation from biogas through steam reforming processes has reported almost the same efficiency as natural gas or methane. An economic investigation by 2013 presented that US$0.27/kWh is attained with payback 8 years (time periods) and some others factors such as maintenance cost, operation and investment (Braga et al. 2013). Since the lower heating value (LHV) of H_2 (120 MJ/kg), the parallelly outcome conclusion of US$8.98/kg$H_2$ (approx. 8 €/ kgH_2) (Braga et al. 2013).

More current results in order to stream reforming connecting membrane reactors confirm that more than 20% increases the system efficiency by considering the membrane reactor. Indeed, the system efficiencies are informed closely to 25–28% along with a fraction of S/C = 3 for a traditional ATR method and between 46–52% and the fraction of S/C = 4 to traditional SR. At the same time, the membrane reactor expresses efficiencies approx.70% along with the fraction of S/C = 3 (Di Marcoberardino et al. 2018).

H_2 production using membrane reactor gets less cost than traditional reforming WGS-PSA methods from economic analysis. In the year 2018, it described that H_2 obtaining 100 kg per day at H_2 cost rate 4 €/kg by using fluidized membrane reactor approach, whereas H_2 cost rates are 4.2 €/kg and 6.4 €/kg respectively described by using the SR and ATR traditional approaches at the same parameters (Di Marcoberardino et al. 2018). The cost of H_2 production from biogas must depend on methane content in biogas—technology scale and carbon price, both factors that are relevant from an economic point of

view. Tradition reforming techniques have been fully established at large scale for reforming natural gas since the globally natural gas generation, which is reached by about 3750 billion m^3 in 2017 and is situated at specific locations (leading to centralized H_2 generation) (Enerdata 2018). Decentralized H_2 generation could be minimum the requirement of distribution infrastructure and reforming methods like dry reforming (DR). Membrane reactors are much more feasible or competitive at a low generation scale. Since biogas has a decentralized feature related to natural gas (CH_4) and a small scale global production of approximately 36 billion m^3 of methane (CH_4) in the year 2015 (Scarlat et al. 2018), it looks more appropriate to decentralized production of H_2. In addition, decentralized H_2 generation makes impractical Carbon Capture and Storage (CCS), which is the method to decrease carbon emissions from the sources of natural gas. Hence, biogas and decentralized generation seem fascinating renewable sources and carbon-free hydrogen production, either if few additions are expensive about impurities removal for biogas reforming like sulphur and siloxanes. Additionally, the dry reforming technique seems to be interesting for biogas reforming and mainly depends on biogas compounds as a reactant.

4. Role Of Operating Parameters

The advancement conversion of biogas needs to address thermodynamic problems such as the need for temperature and high pressure to raise the cost of energy intake. Biogas conversion to syngas requires high temperature (500–800°C) due to endothermic reaction (Chen et al. 2017).

4.1 Reaction Temperature

Reaction temperature has found more influencing parameters on catalyst activity in biogas conversion to syngas. Ni-based catalysts are broadly utilized for biogas conversion to syngas due to their low prices and more potential of catalytic activity than other metals (Zain and Mohamed 2018). An increase in conversion of methane and carbon dioxide rises in temperature attributable to the reaction of endothermicity (Serrano-Lotina and Daza 2014). According to Xu et al. (2010) methane and caron dioxide conversions are higher than the estimation of thermodynamic equilibrium. This could be related to further secondary reactions as a breakdown of methane, steam reforming or RWGS. Shen et al. (2011) informed that renovation of CO_2 is always more than methane conversion due to RWGS reaction. However, Chen et al. (2005) detected higher methane (CH_4) transformation than carbon dioxide (CO_2) transformation, representing that methane decomposition an outsized role on the catalytic performance, as Yasyerli et al. (2011) also conveyed the same process.

Over the catalyst, low variance between the renovation of CH_4 and CO_2 at a temperature less than 450°C or higher than 750°C, RWGS was unfavorable due to endothermic reaction. Conversion of CO_2 was appeared low due to the favorable condition of Boudourad equation. At a temperature of more than 750°C, methane decomposition, methane steam reforming could occur; thus, methane conversion may improve, approaching the transformation of CO_2. A rise in ratio of H_2/CO was detected as the temperature is increased. However, there was below unity in all cases what recommended that reaction of RWGS was

always occur but in low range when temperature is raised (Chen et al.2005; Yasyerli et al.2011). A detail of the effective terms in catalytic conversion of methane to syngas using Ni-based catalysts is offered in Table 3.

Table 3
The impact of operating parameters on methane conversion to syngas at 1 atm.

Catalyst	Temperature(°C)	GHSV(h^{-1})	CH_4 Conversion (%)	CO_2 Conversion (%)	Ref.
3 wt% Gd-promoted Ni/Y_2O_3	700.0	8000	84.0	85.0	(Pechimuthu et al. 2007)
ZrOx/Ni-MnOx/SiO_2	500.0	-	23.0	18.0	(Yao et al. 2017)
15 wt% Ni–ZrO_2	750.0	24000	95.0	98.0	(Zhang et al. 2015)
Yb-promoted Ni (20 wt%)/γ-Al_2O_3	700.0	52000	82.0	89.0	(Amin et al. 2012)
5 wt%Ni – 10 wt%Co/Al_2O_3	700.0	36000	67.0	71.0	(Siang et al. 2018)
15 wt% Ni-MgO-Al_2O_3	750.0	43000	75.0	80.0	(Akbari et al. 2017)
Ni-CeO_2/ZrO_2	600.0	-	55.0	70.0	(Wolfbeisser et al. 2016)
Mg-promoted Co-Ni/SBA-15	700.0	8000	80.0	75.0	(Al-Fatesh et al.2017)
Ni/CeO_2-Al_2O_3	800.0	90000	75.0	85.0	(Fouskas et al. 2014)
10% Ni/γ- Al_2O_3	700.0	1150	90.0	86.0	(Rathod and Bhale 2014)

4.2 Gas Hourly Space Velocity (GHSV)

It is volumetric feed gas flow rate to the volume of the reactor (Gregorie et al. 2020). It is an important parameter to design and optimize purpose in steam reformers and has more effect on conversion on methane distribution of temperature inside the reactor (Sadooghi et al. 2015).

Hamedi et al. (2014) investigated the GHSV ranges from 16,500 per hour to 34,500 per hour and it related to the rate of gas feed flow on reforming performance and temperature profile. An increase in the GHSV changes the location of the maximum temperature drift of the catalyst. However, the extreme temperature remains the same. The normal catalyst temperature is rather higher when raising the flow rate due to the uniform distribution of temperature. It found a low yield of hydrogen to entirely simulated ratios of

O_2/CH_4 when increasing feed gas space velocity. It must be well known that the residence time reduces from 0.22 second to 0.11second.

Avraam et al. (2010) reported catalytic behavior at temperature 715°C of 5%Ru/Al_2O_3 catalyst by the effect of space velocity. Gas feed molar composition was CH_4(25%), CO_2 (25%), H_2O (50%) and 1 bar inlet pressure. The results were found methane conversion raises from 46–93% as GHSV reduces from 76,000–10,000 per hour.

4.3 O_2/CH_4 Ratio

Leading the ratio of O_2/CH_4 increases the peak temperature in the reactor. The concentration of O_2 is increased in feeding gas that promotes the rates of oxidation reaction as additional fuel energy is delivered as a form of heat (Hamedi et al. 2014). Lau et al. (2012) stated that there is no activity of reforming observed at lower exhaust gas temperatures. The main point is that the enthalpy of the reactant is inadequate to control the activation energy barrier of reforming reactions. An increase in the reactor's temperature could be seen by providing the oxygen in reactor feed as a result of incomplete oxidation catalytic incineration of the CH4 nearby entrance of catalyst (when the temperature of the reactor is high than the catalyst temperature light-off) (Lee et al. 2005).

5. Conclusion

Reforming biogas is a promising method to generate green H_2 and minimize the overload on natural gas (CH_4). The main difficulties of biogas reforming methods are raised associated with coke establishment over the catalyst surface done by most researchers and specialists. It also reviewed related to toxic conditions on account of sulfur content which causes catalyst inactivation and diminishes the production of H_2. Ni-based catalysts are mostly used in miscellaneous reforming techniques such as natural gas reforming and biogas reforming. A huge quantity of examining works is also attentive to the effectiveness of catalyst support and some additional promoter elements to reduce the issues of coke formation.

The utilization of catalyst in the reaction of H_2 generation must be advanced to a longer usage life cycle to avoid coke formation over the active catalyst surface. The ideal catalysts are Ni-based in reforming processes because of their lower cost. DR and DOR methods are appeared more attractive techniques for H_2 generation from biogas. However, these methods are not favorable as carbon development on the surface of catalysts and extra oxidant is needed.

ATR process could be attractive because of higher conversion efficiency and H2 yields; however, this technique is a highly complex control system. At present, mostly H_2 generation form biogas is attained in steam reforming method though conventional reformer. However, it is possible to obtain H2 from biogas with advanced membrane reactors (only in one step) and at low temperatures (i.e., reforming process and separation happens at the same location).

GHSV is an essential parameter to design and optimization purpose in steam reformers and has more effect on conversion on methane distribution of temperature inside the reactor. Important operating parameters such as temperature mostly utilized the range from 700-800^{o}C for maximum conversion of CH4 and CO2 and different catalysts as reviewed by this present paper.

The economic analysis point of view that more current results to stream reforming connecting membrane reactors confirm that more than 20% increases the system efficiency by considering the membrane reactor. The cost of H_2 production from biogas must depend on methane content in biogas. Hydrogen generation cost from biogas (0.07 €/ kWh) is higher than natural gas (0.03 €/kWh).

Various reforming methods can be nominated for H_2 production depend on biogas composition, the need for hydrogen purity, investment availability, and quantity of the desired hydrogen.

6. Declarations

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and materials: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Funding: There is no funding.

Competing interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author contribution: Conceptualization: Anil Kumar, and Ravindra Kumar; Formal analysis and investigation: Anil Kumar; Writing - original draft: Ravindra Kumar; Writing - review and editing: Anil Kumar; Supervision: Anil Kumar and Amit Pal.

Acknowledgments: Authors are highly grateful to Department of Mechanical Engineering and Centre for Energy and Environment, Delhi Technological University Delhi (India), for providing basic facilities to compile this work.

7. References

1. Akbari, E., Alavi, S. M., & Rezaei, M. (2017). Synthesis gas production over highly active and stable nanostructured NiMgOAl2O3 catalysts in dry reforming of methane: Effects of Ni contents. Fuel, 194, 171-179. https://doi.org/10.1016/j.fuel.2017.01.018
2. Al-Fatesh, A. S., Arafat, Y., Atia, H., Ibrahim, A. A., Ha, Q. L. M., Schneider, M., & Fakeeha, A. H. (2017). CO_2-reforming of methane to produce syngas over Co-Ni/SBA-15 catalyst: Effect of support

modifiers (Mg, La and Sc) on catalytic stability. Journal of CO_2 Utilization, 21, 395-404. https://doi.org/10.1016/j.jcou.2017.08.001

3. Alves, H. J., Junior, C. B., Niklevicz, R. R., Frigo, E. P., Frigo, M. S., & Coimbra-Araújo, C. H. (2013). Overview of hydrogen production technologies from biogas and the applications in fuel cells. International journal of hydrogen energy,38(13), 5215-5225. https://doi.org/10.1016/j.ijhydene.2013.02.057
4. Amin, M. H., Mantri, K., Newnham, J., Tardio, J., & Bhargava, S. K. (2012). Highly stable ytterbium promoted Ni/γ-Al2O3 catalysts for carbon dioxide reforming of methane. Applied Catalysis B: Environmental, 119, 217-226. https://doi.org/10.1016/j.apcatb.2012.02.039
5. Araki, S., Hino, N., Mori, T., & Hikazudani, S. (2009). Durability of a Ni based monolithic catalyst in the autothermal reforming of biogas. International journal of hydrogen energy, 34(11), 4727-4734. https://doi.org/10.1016/j.ijhydene.2009.03.041
6. Armor, J. N. (1999). The multiple roles for catalysis in the production of H_2. Applied Catalysis A: General, 176(2), 159-176. https://doi.org/10.1016/S0926-860X(98)00244-0
7. Avraam, D. G., Halkides, T. I., Liguras, D. K., Bereketidou, O. A., & Goula, M. A. (2010). An experimental and theoretical approach for the biogas steam reforming reaction. International Journal of Hydrogen Energy, 35(18), 9818-9827. https://doi.org/10.1016/j.ijhydene.2010.05.106
8. Braga, L. B., Silveira, J. L., Da Silva, M. E., Tuna, C. E., Machin, E. B., & Pedroso, D. T. (2013). Hydrogen production by biogas steam reforming: A technical, economic and ecological analysis. Renewable and Sustainable Energy Reviews, 28, 166-173. https://doi.org/10.1016/j.rser.2013.07.060
9. Cai, X., Dong, X., & Lin, W. (2006). Autothermal reforming of methane over Ni catalysts supported on CuO-ZrO_2-CeO_2-Al_2O_3. Journal of Natural Gas Chemistry, 15(2), 122-126. https://doi.org/10.1016/S1003-9953(06)60018-2
10. Chattanathan, S. A., Adhikari, S., McVey, M., & Fasina, O. (2014). Hydrogen production from biogas reforming and the effect of H_2S on CH_4 conversion. International journal of hydrogen energy, 39(35), 19905-19911. https://doi.org/10.1016/j.ijhydene.2014.09.162
11. Chen, L., Zhu, Q., Hao, Z., Zhang, T., & Xie, Z. (2010). Development of a Co–Ni bimetallic aerogel catalyst for hydrogen production via methane oxidative CO_2 reforming in a magnetic assisted fluidized bed. international journal of hydrogen energy, 35(16), 8494-8502. https://doi.org/10.1016/j.ijhydene.2010.06.003
12. Chen, X., Honda, K., & Zhang, Z. G. (2005). A comprehensive comparison of CH_4-CO_2 reforming activities of NiO/Al_2O_3 catalysts under fixed-and fluidized-bed operations. Applied Catalysis A: General, 288(1-2), 86-97. https://doi.org/10.1016/j.apcata.2005.04.037
13. Chen, X., Jiang, J., Li, K., Tian, S., & Yan, F. (2017). Energy-efficient biogas reforming process to produce syngas: The enhanced methane conversion by O_2. Applied Energy, 185, 687-697. https://doi.org/10.1016/j.apenergy.2016.10.114

14. Di Marcoberardino, G., Foresti, S., Binotti, M., & Manzolini, G. (2018). Potentiality of a biogas membrane reformer for decentralized hydrogen production. Chemical Engineering and Processing-Process Intensification, 129, 131-141. https://doi.org/10.1016/j.cep.2018.04.023
15. Enerdata, Global Energy Statistical Yearbook (2018). Accessed 10 June 2019. https://yearbook.enerdata.net.
16. Fouskas, A., Kollia, M., Kambolis, A., Papadopoulou, C., & Matralis, H. (2014). Boron-modified Ni/Al2O3 catalysts for reduced carbon deposition during dry reforming of methane. Applied Catalysis A: General, 474, 125-134. https://doi.org/10.1016/j.apcata.2013.08.016
17. Galvagno, A., Chiodo, V., Urbani, F., & Freni, F. (2013). Biogas as hydrogen source for fuel cell applications. International Journal of Hydrogen Energy, 38(10), 3913-3920. https://doi.org/10.1016/j.ijhydene.2013.01.083
18. Götz, M., Lefebvre, J., Mörs, F., Koch, A. M., Graf, F., Bajohr, S., & Kolb, T. (2016). Renewable Power-to-Gas: A technological and economic review. Renewable energy, 85, 1371-1390. https://doi.org/10.1016/j.renene.2015.07.066
19. Gregorie, E. F., Lamb, J. J., Lien, K. M., Pollet, B. G., & Burheim, O. S. (2020). Hydrogen and biogas. In Micro-Optics and Energy (pp. 131-155). Springer, Cham
20. Hamedi, M. R., Tsolakis, A., & Lau, C. S. (2014). Biogas upgrading for on-board hydrogen production: Reforming process CFD modelling. International journal of hydrogen energy, 39(24), 12532-12540. https://doi.org/10.1016/j.ijhydene.2014.06.017
21. Horikawa, M. S., Rossi, F., Gimenes, M. L., Costa, C. M. M., & Da Silva, M. G. C. (2004). Chemical absorption of H_2S for biogas purification. Brazilian Journal of Chemical Engineering, 21(3), 415-422. https://doi.org/10.1590/S0104-66322004000300006
22. Kalamaras, C. M., & Efstathiou, A. M. (2013). Hydrogen production technologies: current state and future developments. In Conference papers in science (Vol. 2013). Hindawi. http://dx.doi.org/10.1155/2013/690627
23. Kumar, R., Kumar, A., & Pal, A. (2020). An overview of conventional and non-conventional hydrogen production methods. Materials Today: Proceedings. https://doi.org/10.1016/j.matpr.2020.08.793
24. Lau, C. S., Allen, D., Tsolakis, A., Golunski, S. E., & Wyszynski, M. L. (2012). Biogas upgrade to syngas through thermochemical recovery using exhaust gas reforming. Biomass and bioenergy, 40, 86-95. https://doi.org/10.1016/j.biombioe.2012.02.004
25. Lau, C. S., Tsolakis, A., & Wyszynski, M. L. (2011). Biogas upgrade to syn-gas (H_2–CO) via dry and oxidative reforming. International Journal of Hydrogen Energy, 36(1), 397-404. https://doi.org/10.1016/j.ijhydene.2010.09.086
26. Lee, S. H., Applegate, D. V., Ahmed, S., Calderone, S. G., & Harvey, T. L. (2005). Hydrogen from natural gas: part I—autothermal reforming in an integrated fuel processor. International Journal of Hydrogen Energy, 30(8), 829-842. https://doi.org/10.1016/j.ijhydene.2004.09.010
27. Lin, Y. M., Liu, S. L., Chuang, C. H., & Chu, Y. T. (2003). Effect of incipient removal of hydrogen through palladium membrane on the conversion of methane steam reforming: experimental and

modeling. Catalysis Today, 82(1-4), 127-139. https://doi.org/10.1016/S0920-5861(03)00212-8

28. Mahecha-Botero, A., Chen, Z., Grace, J. R., Elnashaie, S. S. E. H., Lim, C. J., Rakib, M., & Shirasaki, Y. (2009). Comparison of fluidized bed flow regimes for steam methane reforming in membrane reactors: A simulation study. Chemical engineering science, 64(16), 3598-3613. https://doi.org/10.1016/j.ces.2009.04.044
29. Minh, D. P., Siang, T. J., Vo, D. V. N., Phan, T. S., Ridart, C., Nzihou, A., & Grouset, D. (2018). Hydrogen production from biogas reforming: An overview of steam reforming, dry reforming, dual reforming, and tri-reforming of methane. Hydrogen Supply Chains, 111-166. https://doi.org/10.1016/B978-0-12-811197-0.00004-X
30. Mosayebi, Z., Rezaei, M., Ravandi, A. B., & Hadian, N. (2012). Autothermal reforming of methane over nickel catalysts supported on nanocrystalline $MgAl_2O_4$ with high surface area. International journal of hydrogen energy, 37(2), 1236-1242. https://doi.org/10.1016/j.ijhydene.2011.09.141
31. Nahar, G., Mote, D., & Dupont, V. (2017). Hydrogen production from reforming of biogas: Review of technological advances and an Indian perspective. Renewable and Sustainable Energy Reviews, 76, 1032-1052. https://doi.org/10.1016/j.rser.2017.02.031
32. Nalbant, Y., & Colpan, C. O. (2020). An Overview of Hydrogen Production from Biogas. Accelerating the Transition to a 100% Renewable Energy Era, 355-373. https://doi.org/10.1007/978-3-030-40738-4_16
33. Nawfal, M. (2015). Valorisation catalytique du biogaz pour une énergie propre et renouvelable (Doctoral dissertation).
34. Nikolaiis, P., & Poullikkas, A. (2017). A comparative overview of hydrogen production processes. Renewable and sustainable energy reviews, 67, 597-611. https://doi.org/10.1016/j.rser.2016.09.044
35. Noh, Y. S., Lee, K. Y., & Moon, D. J. (2019). Hydrogen production by steam reforming of methane over nickel based structured catalysts supported on calcium aluminate modified SiC. International Journal of Hydrogen Energy, 44(38), 21010-21019. https://doi.org/10.1016/j.ijhydene.2019.04.287
36. Ouyang, L. Z., Huang, J. M., Wang, H., Wen, Y. J., Zhang, Q. A., Sun, D. L., & Zhu, M. (2013). Excellent hydrolysis performances of Mg3RE hydrides. International journal of hydrogen energy, 38(7), 2973-2978. https://doi.org/10.1016/j.ijhydene.2012.12.092
37. Patinvoh, R. J., Osadolor, O. A., Chandolias, K., Horváth, I. S., & Taherzadeh, M. J. (2017). Innovative pre-treatment strategies for biogas production. Bioresource technology, 224, 13-24. https://doi.org/10.1016/j.biortech.2016.11.083
38. Pechimuthu, N. A., Pant, K. K., & Dhingra, S. C. (2007). Deactivation studies over Ni− K/CeO_2− Al_2O_3 catalyst for dry reforming of methane. Industrial & engineering chemistry research, 46(6), 1731-1736. https://doi.org/10.1021/ie061389n
39. Rathod, V., & Bhale, P. V. (2014). Experimental investigation on biogas reforming for syngas production over an alumina-based nickel catalyst. Energy Procedia, 54, 236-245. https://doi.org/10.1016/j.egypro.2014.07.267

40. Rau, F., Herrmann, A., Krause, H., Fino, D., & Trimis, D. (2019). Efficiency of a pilot-plant for the autothermal reforming of biogas. International Journal of Hydrogen Energy, 44(35), 19135-19140. https://doi.org/10.1016/j.ijhydene.2018.04.014
41. Sadooghi, P., & Rauch, R. (2015). Experimental and modeling study of hydrogen production from catalytic steam reforming of methane mixture with hydrogen sulfide. International journal of hydrogen energy, 40(33), 10418-10426. https://doi.org/10.1016/j.ijhydene.2015.06.143
42. Sato, T., Suzuki, T., Aketa, M., Ishiyama, Y., Mimura, K., & Itoh, N. (2010). Steam reforming of biogas mixtures with a palladium membrane reactor system. Chemical engineering science, 65(1), 451-457. https://doi.org/10.1016/j.ces.2009.04.013
43. Scarlat, N., Dallemand, J. F., & Fahl, F. (2018). Biogas: Developments and perspectives in Europe. Renewable energy, 129, 457-472. https://doi.org/10.1016/j.renene.2018.03.006
44. Serrano-Lotina, A., & Daza, L. (2014). Influence of the operating parameters over dry reforming of methane to syngas. International Journal of Hydrogen Energy, 39(8), 4089-4094. https://doi.org/10.1016/j.ijhydene.2013.05.135
45. Serrano-Lotina, A., Martin, A. J., Folgado, M. A., & Daza, L. (2012). Dry reforming of methane to syngas over La-promoted hydrotalcite clay-derived catalysts. International Journal of Hydrogen Energy, 37(17), 12342-12350. https://doi.org/10.1016/j.ijhydene.2012.06.041
46. Shanmugam, K., Tysklind, M., & Upadhyayula, V. K. (2018). Use of liquefied biomethane (LBM) as a vehicle fuel for road freight transportation: a case study evaluating environmental performance of using LBM for operation of tractor trailers. Procedia CIRP, 69, 517-522.https://doi.org/10.1016/j.procir.2017.11.133
47. Shen, W., Momoi, H., Komatsubara, K., Saito, T., Yoshida, A., & Naito, S. (2011). Marked role of mesopores for the prevention of sintering and carbon deposition in dry reforming of methane over ordered mesoporous Ni–Mg–Al oxides. Catalysis today, 171(1), 150-155. https://doi.org/10.1016/j.cattod.2011.04.003
48. Siang, T. J., Singh, S., Omoregbe, O., Bach, L. G., Phuc, N. H. H., & Vo, D. V. N. (2018). Hydrogen production from CH_4 dry reforming over bimetallic Ni–Co/Al2O3 catalyst. Journal of the Energy Institute, 91(5), 683-694. https://doi.org/10.1016/j.joei.2017.06.001
49. Sinha, P., & Pandey, A. (2011). An evaluative report and challenges for fermentative biohydrogen production. International Journal of Hydrogen Energy, 36(13), 7460-7478. https://doi.org/10.1016/j.ijhydene.2011.03.077
50. Verma, P., & Samanta, S. K. (2016). Overview of biogas reforming technologies for hydrogen production: advantages and challenges. In Proceedings of the first international conference on recent advances in bioenergy research (pp. 227-243). Springer, New Delhi. https://doi.org/10.1007/978-81-322-2773-1_17
51. Wolfbeisser, A., Sophiphun, O., Bernardi, J., Wittayakun, J., Föttinger, K., & Rupprechter, G. (2016). Methane dry reforming over ceria-zirconia supported Ni catalysts. Catalysis Today, 277, 234-245. https://doi.org/10.1016/j.cattod.2016.04.025

52. Xu, J., Zhou, W., Li, Z., Wang, J., & Ma, J. (2010). Biogas reforming for hydrogen production over a Ni–Co bimetallic catalyst: effect of operating conditions. International Journal of Hydrogen Energy, 35(23), 13013-13020. https://doi.org/10.1016/j.ijhydene.2010.04.075
53. Yao, L., Wang, Y., Shi, J., Xu, H., Shen, W., & Hu, C. (2017). The influence of reduction temperature on the performance of ZrOx/Ni-MnOx/SiO_2 catalyst for low-temperature CO_2 reforming of methane. Catalysis Today, 281, 259-267. https://doi.org/10.1016/j.cattod.2016.05.031
54. Yasyerli, S. E. N. A., Filizgok, S., Arbag, H., Yasyerli, N. A. İ. L., & Dogu, G. (2011). Ru incorporated Ni–MCM-41 mesoporous catalysts for dry reforming of methane: Effects of Mg addition, feed composition and temperature. International Journal of Hydrogen Energy, 36(8), 4863-4874. https://doi.org/10.1016/j.ijhydene.2011.01.120
55. Zain, M. M., & Mohamed, A. R. (2018). An overview on conversion technologies to produce value added products from CH_4 and CO_2 as major biogas constituents. Renewable and Sustainable Energy Reviews, 98, 56-63. https://doi.org/10.1016/j.apenergy.2016.10.114
56. Zhang, X., Zhang, Q., Tsubaki, N., Tan, Y., & Han, Y. (2015). Carbon dioxide reforming of methane over Ni nanoparticles incorporated into mesoporous amorphous ZrO2 matrix. Fuel, 147, 243-252. https://doi.org/10.1016/j.fuel.2015.01.076

Figures

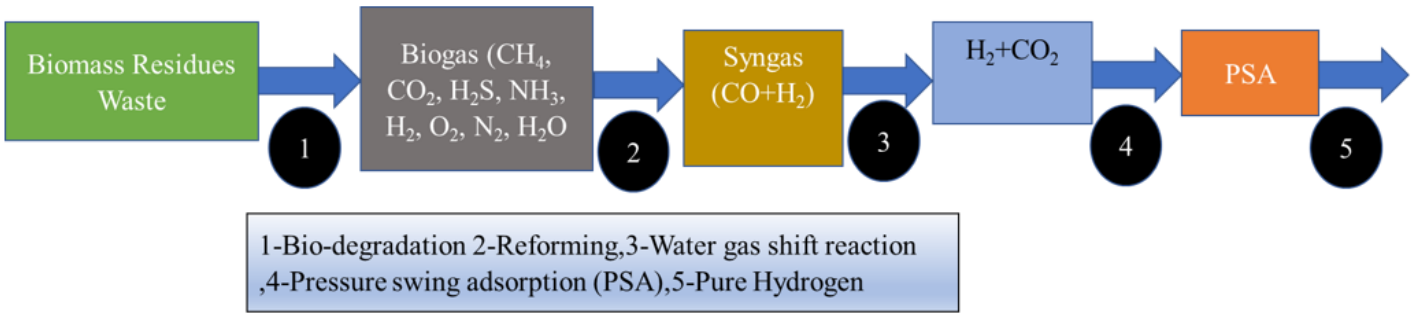


Figure 1

The flow process of hydrogen generation from biogas (Nalbant and Colpan 2020)

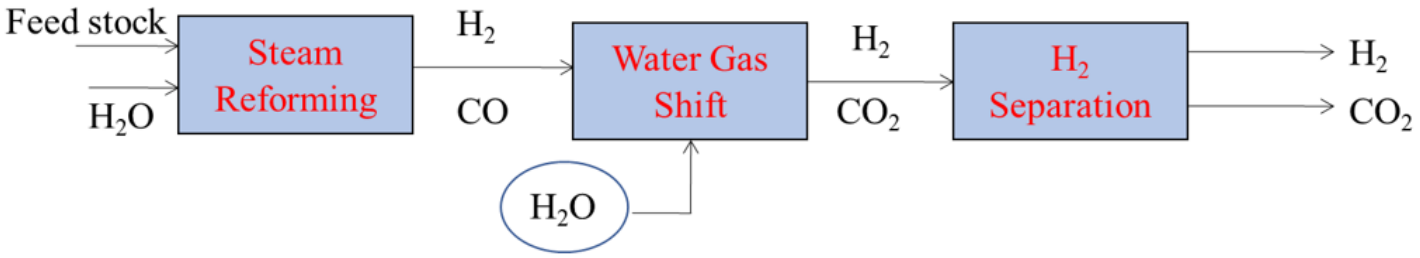


Figure 2

Flow process of steam reforming process for generation of hydrogen (Kumar et al. 2020)

IAF Transition towards a Network-Centric-Knowledge Based Force

Author: Wing Commander Mukesh Kumar Singh & Dr Vikas Gupta

Period: January 2021 - March 2021

Abstract

The article is based upon a study done by the authors regarding Knowledge management in the Indian Air Force (IAF). The study brings out valuable lessons regarding institutionalising the way to retain the knowledge base of the organisation keeping in mind that personnel moving out of the organisation is inevitable. Though the study is based upon the IAF the findings are of equal relevance to the complete military.

Introduction

Knowledge Management (KM) is getting the right knowledge to the right person at the right time. In a military environment, KM includes a strategic approach to achieve strategic aims by using the power of collective knowledge penetration into the processes of generating, gathering, organising, sharing and transferring knowledge to meet the military objective.1

Knowledge is power, therefore an asset for people and organisations. It has superseded traditional corporate assets and has become a strategic resource that drives the economy and provides a lasting competitive advantage2,3. Leveraging individual and organisational knowledge to boost efficiency and effectiveness has now become a popular management technique that has been successful across the globe. A survey conducted in more than 200 successful US firms revealed that 80% of corporations had knowledge initiatives. There is adequate evidence in numerous studies conducted worldwide that effective KM is beneficial to the organisation in saving cost and time, increasing innovations, efficiency, customer satisfaction, competitive edge, and organisational performance4,5,6,7,8.

Like any other organisation, the requirement of an effective KM in a contemporary military cannot be ruled out. Studies indicate that with the expansion of operational areas and sub-conventional warfare taking the center-stage, a large number of military organisations like NATO, US, UK, Israel, Canada, Turkey etc. have adopted KM in a big way. The cutting-edge technologies have resulted in dynamic, unpredictable and complex operations. This makes

problem-solving and decision making more complex and essential than ever. These significant changes in the characteristics of warfare dynamics have forced the militaries in the world to reorganise their doctrines to include KM at the core9.

Aggressive KM initiatives also assumes a vital role for the IAF as timely, accurate, decision-ready and actionable knowledge is imperative for planning and conducting aerospace operations10. However, despite the best efforts, IAF continues to squander what may be its greatest assets in today's knowledge economy i.e. the wealth of experience, ideas, and insight of its air-warriors that are scattered across and deeply embedded within the organisation11. The IAF is required to move beyond the simplistic incorporation of the latest information technology (IT) hardware and software to a more deliberate KM strategy. There is a need to progress beyond the simple network-centric force to a network-centric-knowledge-enabled force to achieve knowledge superiority.

KM in the World's Military Organisations

The available literature establishes KM as a force multiplier for the military establishments. Hence, in the 21st-century, military organisations of the world have no option other than to take cognisance of the fact and embrace KM in a big way to remain current and relevant. Some of the contemporary military establishments have already initiated necessary steps to incorporate KM in their doctrines and policies12,13.

The NATO Knowledge vision declares that "the NATO military structure will transform into Knowledge Centric Organisation. It will promote information and knowledge sharing and treat expertise, experience and best practices as valuable assets to achieve decision superiority"14. The United Nations (UN) peacekeeping operations have developed a "Knowledge-Based System" to assist the UN Military Observers in the deployment country15.

The US Department of Defence directs the Joint Forces to be knowledge empowered and initiate actions to acquire, refine and share knowledge. KM was included in the US Army Regulations to emphasise KM as the means of supporting its strategy for the 21st century, which is of transforming itself into a net-centric, knowledge-based force. Similarly, the US Air Force Strategic Plan enforces Knowledge-Based Operations, which envisages effective use and management of information and knowledge across the Air Force16. The Israeli Defence Forces foresee knowledge as critical resources for current and future battlefield17. The Defence Research and Development Canada has also targeted KM as an area to support the futuristic military requirements in 21st century18. The Armed Forces of the Slovak Republic also believes in active KM for retention of necessary knowledge19. Thus, KM is being given importance by several militaries.

KM Status Check of the IAF

Two of the major impediments faced by the military organisations in managing the organisational knowledge is the consistent knowledge loss because of soldiers' exit and lack of formalised KM strategy20. Studies show that half or more of the best officers leave the military early, rather than serving for a full career21. Also, to keep their forces young and fighting fit, the military organisations by the policy itself have an option for voluntary retirement after 15 to 20 years of service. As per the statistics, 89% of the soldiers of all the armed forces retire young between 30–50 years of age22. To examine the status of IAF on these two aspects, a survey study was done to address the following questions:

Q1. The effects on the organisation's knowledge due to the early retirement of its air-warriors?

Q2. Status of existing KM strategies and plans in IAF?

Study Methodology

An inferential, descriptive study was employed to examine the status of KM in IAF. This was a quantitative study in nature, utilising descriptive statistics to analyse the data derived from the KM survey. A survey instrument was designed to gather data from 250 retired air-veterans, comprising of 100 officers and 150 airmen. Since the complete population of air-veterans could not be ascertained, snowball sampling was used to select the participants for this study wherein each sample was requested to refer a friend or acquaintance air-veteran for participation. The survey questionnaire was divided into two sections, comprising of four questions in each. Respondents were asked to record the answers on a 5-point Likert scale ranging from (1) strongly agree, with (3) neutral as the midpoint, and (5) strongly disagree. Pilot testing was conducted to establish content validity and improve questions. Cronbach's Alpha reliability coefficient23 was used to measure the internal consistency of the questions. The questions were found to be reliable with a 0.773 Cronbach Alpha coefficient for all the eight items. This study used SPSS 23.024 as major tools to analyse data.

Results and Analysis

To better understand the characteristics of each variable, descriptive statistics analysis was used. The questions and responses are as under:

Q 1: What are the effects of knowledge loss in IAF due to the early retirement of air-warriors?

Response of sample population on various questions asked to understand the effects on the organisation's knowledge due to the early retirement of the air-warriors indicates:

(a) 93.2% of the respondents feel that the retiring air-warriors possess critical professional knowledge, the loss of which adversely affects the organisation's performance.

(b) 51.6 % feel that that the existing system and process in IAF caters to redundancies with adequate induction and training, so that early retirement does not have any significant effect on the organisations knowledge, while 34.4% don't agree to this fact.

(c) 84% of the respondents believe that early retirement by 20 yr of service, creates a vacuum at the senior supervisory level in a technology-intensive organisation.

(d) 50.4% of the respondents feel that early retirement is good for the organisation as it helps to maintain a young and fit workforce with greater adaptability and current knowledge, whereas 28% disagree with the facts and the remaining
21.6% remained neutral on this question.

Q2: What is the status of the existing KM strategies in the IAF?

Response of sample population on various questions asked to understand existing KM strategies and plans in IAF indicates:

(a) 51.6% of respondents believe that the organisation has comprehensive and documented KM strategies, while 30.4% disagree with the fact and the rest remained neutral.

(b) 45.6% of the respondents feel that the KM roles and responsibilities are clearly defined in the organisation while 34.8% expressed disagreement.

(c) Mixed opinion was received on whether the HR policies in the organisation support the proper management of individual and organisational knowledge. 42% agreed to the fact whereas 40.4% expressed disagreement.

(d) 71.6% of the respondents strongly believe that KM is an integrated part of routine work processes, but it is practiced without realising it.

Average Mean from Officers and Airmen indicated that Officers are more in disagreement on the existence of a formalised KM strategy in IAF as compared to Airmen.

Key Findings

- Retiring air-warriors possess critical professional knowledge, the loss of which adversely affects the organisation's performance. It also creates a vacuum at the senior supervisory level. However, early retirement cannot be completely avoided as it helps to maintain a young and fit workforce. Therefore, the only viable option available is to capture and transfer the knowledge and expertise of these air-warriors.
- KM is an integrated part of routine work processes in IAF, however the same is neither formalised nor documented and practiced sparingly without realising it.
- Officers strongly feel a lack of KM strategy and plans in IAF as compared to Airmen.

Recommendations

- Organisational and individual knowledge be treated as assets and it needs to be managed to attain knowledge superiority.

- The relevance of active KM needs to be acknowledged at the highest levels and driven down to all ranks through Policies, Doctrines and defined Roles and Responsibilities.
- Develop a KM strategy for the IAF that is closely aligned with the IAF's overall strategy and goals.
- Capturing and Transferring of knowledge needs to be emphasised and methods may be instituted to incorporate the same in the routine process.
- Knowledge audit may be conducted to understand the organisational knowledge needs, holdings and deficiencies.
- Efficacy of various KM Tools like Community of Practices, Lesson learned, Knowledge Portal etc. for the IAF and need to be implemented in a planned manner.

Conclusion

The departing soldiers at the peak of their ability and experiences take with them the most valuable knowledge they have gained over the years of their professional and personal lives. Their in-depth knowledge experience, aptitude to understand and integrate knowledge reduces the organisation's capability of innovating and problem-solving which invariably decreases the organisation's performance and profitability. It is in the strong interest of IAF to identify optimal strategies to leverage the intellectual capital of the organisation's workforce. The IAF, therefore is required to develop a KM strategy, to transform from information-base to knowledge-based operation and turn into a net-centric-knowledge-based force.

Endnotes

1 Rahmati, Dariush, Khosro Eskandari, Eghbal Hosseini Sadr, et Morad Ali Nouri. *Providing an Applied Model for Knowledge Management Development Planning in Military Organizations: Providing Proposed Methodology for Naja. [En. (Kuwait Chapter of the Arabian Journal of Business and Management Review* 4, no 1, 2014): 532.

2 Gupta, Vikas, et Meenu Chopra. *Gauging the impact of knowledge management practices on organizational performance–a balanced scorecard perspective. [En. (VINE Journal of Information and Knowledge Management Systems* 48, no 1, 2018): 21-46.

3 Gupta, Vikas, et Namita Jain. *Harnessing information and communication technologies for effective knowledge creation: Shaping the future of education. [En. (Journal of Enterprise Information Management* 30, no 5, 2017): 831-55.

4 Zaied, Abdel Nasser H, Gawaher Soliman Hussein, et Mohamed M Hassan. *The role of knowledge management in enhancing organizational performance. [En. (International Journal of Information Engineering and Electronic Business* 4, no 5, 2012): 27.

5 McQuade, Eamonn, Ellen Sjoer, Peter Fabian, José Carlos Nascimento, et Sanaz Schroeder. *Will you miss me when I'm gone? A study of the potential loss of company knowledge and expertise as employees retire. [En. (Journal of European industrial training* 31, no 9, 2007): 758-68.

6 Sinha, Neena, NK Kakkar, et Vikas Gupta. *Unleash the power of creativity and innovation. [En. (International Journal of Sustainable Strategic Management* 1, no 4, 2009): 417-33.

7 Lin, Tung-Ching, Christina Ling-hsing Chang, et Wen-Chin Tsai. *The influences of knowledge loss and knowledge retention mechanisms on the absorptive capacity and performance of a MIS department.* [En. (Management Decision 54, no 7, 2016): 1757-87.

8 Massingham, Peter Rex. *Measuring the impact of knowledge loss: a longitudinal study. [En. (Journal of Knowledge Management* 22, no 4, 2018): 721-58.

9 Orhan, SE Sensoy A Keskin ME. *Knowledge Management in Military Organizations: Applications of Knowledge Creation and Knowledge Transfer. [En. (Military and Security Studies 2015,* 2015): 20.

10 Air Force Instructions (2014).

11 DiGiacomo, Joseph. *Implementing knowledge management as a strategic initiative.* Naval Postgraduate School Monterey CA (2003).

12 Sullivan, Patrick K. *The Difference that Makes a Difference: Distinguishing Between Knowledge Management and Information Management in the US Army.* Army Command and General Staff Coll Fort Leavenworth KS School of Advanced … (2011).

13 Force, Indian Air. « Basic Doctrine of the Indian Air Force. »: Citeseer, 2012.

14 Jabloński, Jaroslaw, et Andrzej Lis. *Lessons Learned System as a tool of Managing Organisational Knowledge: The Case of Military Organisations. [En. (Business and Non-Profit Organizations Facing Increased Competition and Growing Customers' Demands* 18, 2012): 163.

15 Marzukhi, Syahaneim, Nur Hidayah Mohammad Daud, Zuraini Zainol, et Omar Zakaria. « Framework of Knowledge-Based System for United Nations Peacekeeping Operations Using Data Mining Technique. » Présentation faite à la 2018 Fourth International Conference on Information Retrieval and Knowledge Management (CAMP), 2018.

16 Palos, Guillermo A. *Communities of Practice: Towards leveraging knowledge in the military.* Naval Postgraduate School Monterey CA (2007).

17 Ariely, Gil. *Operational knowledge management in the military. Dans Knowledge Management: Concepts, Methodologies, Tools, and Applications,* 2376-86: IGI Global, 2008.

18 McIntyre, SG. *Knowledge management in Defence R&D Canada.* Defence Research And Development Canadaottawa (Ontario) (2002).

19 Petrufova, Maria. *Strategic Knowledge Management in The Armed Forces Academy Of General Mr Štefánik In Liptovský Mikuláš, The Slovak Republic. [En. (Journal of Defense*

Resources Management (JoDRM) 6, no 2, 2015): 51-56.

20 Lin, Tung-Ching, Christina Ling-hsing Chang, et Wen-Chin Tsai, op cit.

21 Sofat, Pradeep. *HRM in the Armed Forces Transition of a Soldier to a Second Career.* New Delhi – 110 010 2016. https://idsa.in/monograph.

22 Ibid.

23 Cronbach's alpha is a coefficient of reliability (or consistency).

24 SPSS® Statistics is the world's leading statistical software that is used to solve business and research problems.

Wing Commander Mukesh Kumar Singh is an Aeronautical Engineer in the IAF. He is B.E in E&E, MBA in Information Technology and M.Tech from IIT, Roorkee. He is also pursuing Ph.D from Delhi School of Management.

Dr Vikas Gupta is an Assistant Professor, Delhi School of Management, Delhi Technological University (Former Delhi College of Engineering). He has many research papers and articles in his credit which are published in International and Indian journals. His research interests are in knowledge management, innovation, and corporate social responsibility.

Journal of the United Service Institution of India, Vol. CLI, No. 623, January-March 2021.

Applied Intelligence
https://doi.org/10.1007/s10489-021-02403-5

Identifying influential nodes in weighted complex networks using an improved WVoteRank approach

Sanjay Kumar[1] · **Ankit Panda**[2]

Accepted: 30 March 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Influence maximization is an important research problem in the field of network science because of its business value. It requires the strategic selection of seed nodes called "influential nodes," such that information originating from these nodes can reach numerous nodes in the network. Many real-world networks, such as transportation, communication, and social networks, are weighted networks. Influence maximization in a weighted network is more challenging compared to that in an unweighted network. Many methods, such as weighted degree rank, weighted h-index, weighted betweenness, and weighted VoteRank techniques, have been used to order the nodes based on their spreading capabilities in weighted networks. The VoteRank method is a popular method for finding influential nodes in an unweighted network using the idea of a voting scheme. Recently, the WVoteRank method was proposed to find the seed nodes; it extends the idea of the VoteRank method by considering the edge weights. This method considers only 1-hop neighbors to calculate the voting score of every node. In this study, we propose an improved WVoteRank method based on an extended neighborhood concept, which takes the 1-hop neighbors as well as 2-hop neighbors into account for the voting process to decide influential nodes in a weighted network. We also extend our proposed approach to unweighted networks. We compare the performance of the proposed improved WVoteRank method against the popular centrality measures, weighted degree, weighted closeness, weighted betweenness, weighted h-index, and weighted VoteRank on several real-life and synthetic datasets of diverse sizes and properties. We utilize the widely used stochastic susceptible–infected–recovered information diffusion model to calculate the infection scale, the final infected scale as a function of time, and the average distance between spreaders. The simulation results reveal that the proposed method, improved WVoteRank, considerably outperforms the other methods described above, including the recent WVoteRank.

Keywords Complex Network · Influence maximization · Node centrality · SIR model · Social network · Weighted networks · WVoteRank

1 Introduction

Numerous real-world networks such as air traffic networks, road networks, email networks, and collaboration networks, social network are complex and weighted networks [1, 2].

✉ Sanjay Kumar
sanjay.kumar@dtu.ac.in

Ankit Panda
ankitsp02@gmail.com

[1] Department of Computer Science and Engineering, Delhi Technological University, Main Bawana Road, New Delhi, 110042, India

[2] Department of Computer Science and Engineering, Bennett University, Greater Noida, UP 201310, India

A weighted complex network can be represented as a graph $G = (V, E, W)$, where V represents a set of people who are members of the network, E denotes edges, and W is a weight function from E to the set of real numbers R. In case of social networks, there is an edge between two nodes if they are related by a relation, such as friendship, follower, or followee. Similarly, in case of airport networks, the edge between two nodes or cities represent the number of direct flights between two cities. The weight of an edge between two nodes usually depends on the exchange of services, intensity, or duration [3]. If two nodes frequently interact or if they have numerous interactions, then information is more likely to be transferred between them [4]. The degree of interaction between nodes can greatly affect the information diffusion between them. Influential nodes can effectively propagate the information to a large extent in

Published online: 31 May 2021

complex networks by employing word-of-mouth strategy and mutual trust between the network users. The real-world complex networks are vast and everyday evolving and require special methods to find the set of the most influential nodes in the network, which can be used for purposes such as viral marketing and influence maximization [5, 6]. Influence maximization is of great value because of the benefit of the business and commercial exploitation of the products, which dictate the choices of source spreaders to maximize the final extent of information propagation in the network. Mathematically, influence maximization (IM) problem is expressed as follows: "Given a complex network G, information propagation model M, and a positive integer k, the objective is to choose a set S^* of k users from V as the seed set to maximize the information spread."

$$\sigma_{G,M}(S^*) = \arg \max_{|S|=k \wedge S \subseteq V} \sigma_{G,M}(S) \tag{1}$$

where $\sigma_{G,M}(S^*)$ is the expected numbers of users infected by S^*.

Therefore, choosing efficient spreader source nodes is essential for the maximum spread of information. Members of the social network who are capable of influencing numerous other network members are chosen as the seed nodes to spread information. The process of influence maximization consists of two activities: identification of the source spreader nodes and the information diffusion phase. The propagation of information in a complex network is somewhat identical to epidemic spreading [7–9], and many popular methods for modeling information diffusion are based on epidemics spreading. The topological position of nodes and their connectivity with other nodes intuitively play an essential role in determining their spreading capabilities [10, 11]. Based on the notion of the topological significance of nodes, many node centrality methods for unweighted networks have been proposed by researchers. Degree centrality [10] is one of the most commonly used centrality measures to measure the importance of a node in a network, and it is equal to the number of direct neighbors of the node. Degree centrality is a local property of a node in a network. Degree centrality is easily extended in a weighted network [12], and, accordingly, the strength of a node (i) is equal to the sum of edge weights of the edges incident on node i, as given in the following equation:

$$C_s(i) = \sum_{j \in \tau(i)} w_{ij} \tag{2}$$

where $\tau(i)$ is a set containing adjacent neighbors of node i and w_{ij} is the edge weight between node i and j. The closeness and betweenness of the nodes depend on finding the shortest paths between nodes and their costs [13, 14]. The closeness centrality of a node (i) depends on the path lengths between node i to all other nodes in the network. It is equal to the inverse of the sum of the edge weights between node i and all other nodes:

$$C_c(i) = \frac{1}{\sum_{j \in \tau(i)} w_{ij}} \tag{3}$$

The betweenness centrality (C_b) of node i is the fraction of shortest paths passing through that node and is expressed as follows:

$$C_b(i) = \frac{p_{jk}(i)}{p_{jk}} \tag{4}$$

where p_{jk} denotes the number of shortest paths between pair of nodes j and k and $p_{jk}(i)$ denotes such paths that pass through the node i.

Hirsch [15] initially formulated the h-index as a quantitative measure to predict the impact of a scientist's contributions in the field of research. In a social network, the h-index centrality of a node is the maximal value h such that it has at least h neighbors, each of degree h or more. Hence, the h-index is a widely used measure for estimating the influencing power of a node in a network. Furthermore, it has also been proven that the degree, h-index, and coreness of a vertex are related by an operator H [16]. Yu et al. [17] extended the idea of the h-index method on an unweighted network to a weighted network by assuming the edge weights as the order to quantify the spreading ability of links. They first defined the weight of edge w_{ij} as

$$w_{ij} = k_i * k_j \tag{5}$$

where k_i and k_j are the degrees of nodes i and j, respectively, and proposed weighted h-index of a node as the sum of the weighted h-indexes of all neighbors.

$$Whi = \sum_{j \in N(i)} Wh_j \tag{6}$$

where $N(i)$ is the set of neighbor nodes of node i. The weighted h-index method performs better than many classical methods for spreading information.

Finding influential nodes in weighted complex networks to achieve influence maximization is more challenging then unweighted ntworks due to the addition of edge weights. In case of weighted networks edge weight can have different interpretation like capacity or capability of the tie, exchange of services, intensity, or duration of interaction between two nodes. A limited number of research work have been done towards finding influential nodes in weighted nework as cmpared to unweighted netwprks. In this study, we introduce an improved weighted VoteRank method for spreader selection in a weighted network by considering both direct and 2-hop distance neighbors for the voting process. In each iteration, our method selects a top node based on the results of the voting process. The proposed algorithm is a semi-local method intends to properly covers the neighborhood of a node ubder

consideration, leading to an overall high influnece spread. We peform the investigation of our method on many real-life complex networks of different sizes and applications using susceptible–infected–recovered information diffusion model. The main contributions of this study are as follows:

- We improve the performance of the recently introduced WVoteRank method for weighted networks by using the idea of an extended neighborhood. The proposed method ensures the selection of spreader nodes from a diverse position for influence maximization.
- We examine and evaluate conventional methods along with the proposed approach on numerous real-world datasets.
- We determine that the proposed method has a linear running time in terms of input and outperforms many existing methods in terms of different performance measures.

The rest of the paper is organized as follows: In Section 2, we discuss some recent techniques related to influence maximization. In Section 3, we propose an improved weighted VoteRank algorithm based on the extended neighborhood concept to identify influential nodes in weighted complex networks. We also discuss an algorithm for unweighted complex networks and analyze the time complexity of our algorithm. In Section 4, we discuss performance measures for evaluating algorithms for identifying influential nodes in complex networks. A small toy network and the datasets used in this study are presented in Section 5. Section 6 presents an analysis of the results of the proposed improved weighted VoteRank algorithm against some of the various existing algorithms including the WVoteRank method on eight real-life datasets using different performance matrices. Finally, we summarize and conclude the paper in Section 7.

2 Related work

Most contemporary methods used to identify various influential spreaders are designed for unweighted networks. A method for finding influential nodes on a weighted network must consider both the weight and degree of a node. Hence, finding influential nodes on a weighted network is more complicated than that in an unweighted network because of the edge weights between pairs of nodes. Conventionally, it has been observed that a node present in the core of the network has a greater spreading capability, and this leads to the K-shell centrality proposed by Kitsak et al. [18]. The K-shell method continuously removes the lesser degree nodes from the network. This process hierarchically prunes all the nodes until no nodes are present in the network [19]. A node with a higher K-shell value means that it is more central than a node with a lower K-shell value. However, the nodes present in the same shell have the same K-shell values. Bae et al. [20] extended the idea of the K-shell to distinguish between nodes that are in the same K-shell, and they proposed neighborhood coreness (NC) for the unweighted network as follows:

$$NC(v) = \sum_{u \in N(v)} Ks(u) \tag{7}$$

where $Ks(u)$ and $NC(v)$ denote k-shell value and neighborhood coreness of a node u, respectively. Garas et al. [21] generalized the K-shell decomposition method on a weighted network by taking both the weight and degree of the network into account. The method assigns the same shell value as that captured by the classical K-shell decomposition after the removal of edge weights. Yang et al. [22] introduced remaining minimum degree (RMD) decomposition by eliminating the nodes with a minimum degree subsequently. Based on RMD decomposition, a weighted degree (WD) is developed by RMD indices of the nearest neighbors of a node.

Wang et al. [23] introduced an influence spreading model for unweighted networks using the idea of fluid dynamics by depicting the influence spreading method as a fluid update method. Berahmand et al. [24] proposed a local centrality-based spreading method that ranks a node based on its degree, the degree of its neighbors, common connections between the neighbors of the node, and an inverse local cluster coefficient. Souza et al. [25] devised network seeding under budget constraints useful for unweighted networks. They emphasized relatively inexpensive seed nodes neighboring expensive central nodes by following a deterministic fixed threshold information diffusion approach. In a network, community structure [26] can be obtained by finding a group of nodes sharing some commonalities. A community in a social network consists of people who are more closely related to one another compared to other people of different communities. In a community, information originating from a member can reach all other members promptly. Huang et al. [27] recommended a heuristic influence maximization technique applicable for unweighted networks. Their technique combines community detection and topic awareness into influence diffusion modeling. We present some related methods for the influence maximization that inspired us for this research work.

2.1 VoteRank

Zhang et al. [28] proposed the VoteRank method to determine influential nodes in an unweighted network using the idea of the voting process. A tuple (S_v, Va_v) is assigned to each node $v \in V$, where S_v and Va_v correspond to the

vote score and the voting ability of node v, respectively. The vote score S_v of a node (v) is equal to the sum of the voting ability of all its direct neighbors: , i.e.

$$S_v = \sum_{i \in \tau(v)} Va_i \tag{8}$$

where $\tau(v)$ is the set of neighbors of v. The VoteRank method is composed of four phases:

(i) Initialization: The voting score (S_v) of each vertex v is initialized to 0 and the voting ability (Va_v) is initialized to 1.
(ii) Voting: In this phase, voting is conducted. Every vertex v obtains a vote, which is equal to the sum of the voting abilities of its direct neighbors. The vertex that obtains the highest vote is elected as a spreader node. Furthermore, the selected node does not participate in the voting process in the next iteration onward by having its voting score set as 0.
(iii) Update: To find the seed nodes from the distinct locations, a reduction is made in the voting score of the neighboring nodes of a selected node in the subsequent iteration by a value δ=1/$\langle k \rangle$, where $\langle k \rangle$ is the average node degree.
(iv) Iteration: Steps (ii) and (iii) are iterated until c influential nodes are selected, where c is a constant.

VoteRank centrality is applicable for an unweighted network and exhibits much better performance than that of some classical methods. The main benefit of the VoteRank method is that spreaders are selected from diverse locations.

2.2 WVoteRank

Sun et al. [29] extended the concept of the VoteRank approach for the unweighted network to WVoteRank for weighted networks by considering the number of neighbors and the edge weight. Numerous real-life networks are weighted networks, in which the cost of an edge expresses the distance or degree of communication between connecting vertices. By using the number of neighbors and the edge weight, the WVoteRank method computes the voting score S_v of node v as follows:

$$S_v = \sqrt{|\tau(v)|(\sum_{i \in \tau(v)} Va_i \times w(v,i))} \tag{9}$$

where $|\tau(v)|$ is the number of neighbors of node v, $w(v,i)$ is the edge cost between node v and its neighbor i. The selection of spreaders and update phase of the WVoteRank approach is identical to the VoteRank method.

2.3 NCVoteRank

In the VoteRank approach, one assumes that each node has the same voting capability. However, the voting ability of the nodes should be distinct, and it may depend on the topological location of the node in the network. Based on this notion, we [30] previously assumed the neighborhood coreness value of the node as its voting ability and proposed the NCVoteRank method. In this method, the voting score of a node (v) is computed from the following equation:

$$S_v = \sum_{i \in \tau(v)} (Va_i \times NC(i) \times (1-\theta) + Va_i \times \theta) \tag{10}$$

where θ is a regulating parameter and its value ranges between 0 and 1, and $NC(i)$ is the normalized neighborhood coreness of node i, where node i is a direct neighbor of node v. This method performs well, but it is applicable only to unweighted networks.

Note that WVoteRank differs from VoteRank in the calculation of S_v only, where weights of edges are used, and hence it can handle weighted networks. The value of $w(v,i)$ is 1 for unweighted networks, and the formula for S_v in WVoteRank (9) reduces to the formula for S_v in VoteRank. Therefore, WVoteRank is an extension of VoteRank to weighted networks. NCVoteRank differs from VoteRank in the calculation of S_v only. If the value of the controlling parameter (θ) is 1 (10), then NCVoteRank reduces to VoteRank. Therefore, NCVoteRank is also an improvement of VoteRank as one can control θ to achieve better performance. However, like VoteRank, NCVoteRank cannot handle weighted networks. All three techniques, namely, VoteRank, WVoteRank, and NCVoteRank, consider only the 1-hop (direct) neighbors of a node for the voting process to decide the influential node in each iteration. In the following, the proposed improved WVoteRank algorithm considers 1-hop neighbors as well as 2-hop neighbors for the voting process, which adequately covers the neighborhood, leading to an overall better selection of seed nodes.

3 Proposed algorithm: Improved WVoteRank

In this section, we propose an improved WVoteRank method to identify influential spreaders in both weighted and unweighted networks. The proposed method is an extension of the WVoteRank [29] method. The WVoteRank approach takes only direct neighbors or 1-hop neighbors into consideration for the voting process to select the influencing nodes in a network. However, nodes other than direct neighbors can also play an essential role in the selection of source spreaders. Our method considers the 1-hop neighbors as well as 2-hop neighbors in the voting

process. Furthermore, we make the necessary changes in the voting and update phases of the method. We also extend the proposed method to unweighted networks. A tuple (S_v, Va_v) is associated with each node $v \in V$, where S_v and Va_v signify the voting score and voting ability of node v, respectively. The voting score (S_v) is obtained from neighbor nodes up to two hops and is calculated using (5), i.e., $S_v = \sum_{i \in \tau(v)} Va_i$, where $\tau(v)$ corresponds to the adjacent node of v up to two hops.

The workflow of our proposed method is depicted in Fig. 1. It consists of two parts: selection of seed nodes and performance evaluation. The first part finds the constant number of seed nodes (c), and the second part evaluates the proposed method based on various performance parameters. The detailed description of the seed node selection of the proposed work, comprising four phases, is as follows:

(i) Initialization: For each node (v), assign voting score (S_v) as 0 and voting ability (Va_v) as 1.

(ii) Voting: As nodes other than the direct neighbors in the system may play a crucial role in the selection of source spreaders, we consider neighbor nodes up to two hops as contributing in the voting to elect the source spreader. Hence, each vertex v acquires votes equal to the sum of the voting abilities of its direct as well as indirect neighbors by the following equation:

$$S_v = \sqrt{|\tau(v)|(\sum_{i \in \tau(v)} Va_i * w(v,i)) + \sum_{j \in \tau(i)} Va_j * w(i,j)}, \tag{11}$$

where node i denotes the direct neighbor of node v, $w(v,i)$ is the edge weight between node v and node i, $\tau(v)$ is the number of direct neighbors of node v, and j denotes the direct neighbor of node i. Here, node j denotes a node that is two hops away from node v. From (7), it is evident that the voting score of a node v depends on two things: (a) the number of direct neighbors ($\tau(v)$) of node v, voting ability (Va_i) of each direct neighbor i, and the edge weight ($w(v,i)$) between node v and node i and (b) the voting ability (Va_j) of each neighbor j at two hops from node v and the edge weight ($w(i,j)$ between node i and node j.

In the case of an unweighted network, the value of each edge weight $w(v,i)$ is taken as 1, and the vote score obtained by node v is computed as follows:

$$S_v = \sqrt{|N(v)|(\sum_{i \in \tau(v)} Va_i) + \sum_{j \in \tau(i)} Va_j}. \tag{12}$$

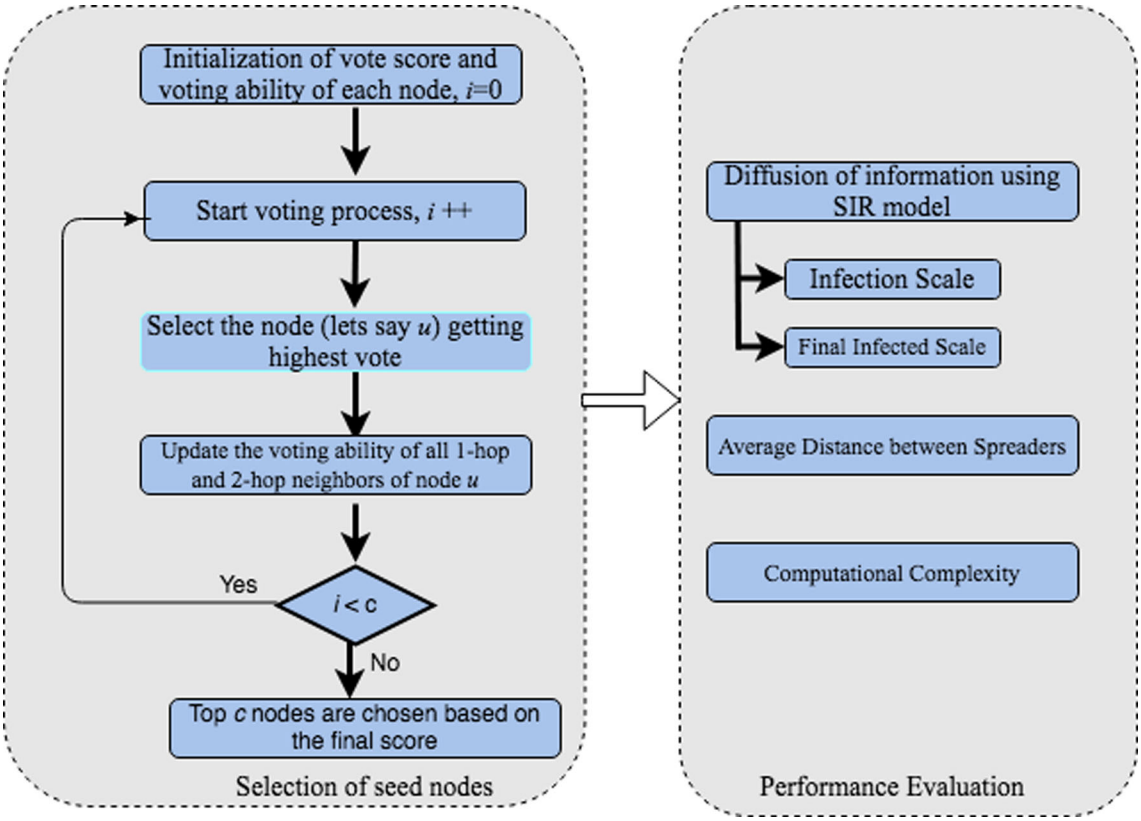


Fig. 1 Major workflow of the proposed algorithm

The node that achieves maximum number of vote is elected as an influential node in the current round if this node was not elected earlier. In subsequent iterations , the same node will not participate in the voting process, and its voting ability is set to 0.

(iii) Update: As discussed in the voting phase, for a particular node u, all its neighbors up to two hops are involved in voting. Therefore, when a node is elected as a spreader, it can influence its neighbors up to two hops. The voting ability of a node that has voted and is up to two hops from the selected spreader is reduced by a factor of δ. The following equation indicates the updated voting ability of the neighbor nodes:

$$Va_v = \begin{cases} Va_v - \delta & \text{if } Va_v - \delta > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{13}$$

where

$$\delta = \frac{1}{kd}$$

and k is the average degree of the nodes, $d = 1$ for the direct neighbor nodes, and $d = 2$ for the nodes that are at two hops from the selected spreader. Therefore, once a node is elected as a spreader, the voting abilities of all adjacent nodes up to two hops are changed using (13). This reduction is made for all neighbors up to two hops. The notion behind the above reduction (δ) is to ensure that, when a node is chosen as a spreader, its neighbor up to two hops has a lower chance of being selected as a spreader in the subsequent round of the voting phases. This results in the selection of influential nodes from diverse positions in the network, leading to overall high information spreading.

(iv) Iteration: Steps (ii) and (iii) are iterated until c influential nodes are selected, where c is a constant.

It is to be noted that the three related methods, namely, VoteRank, WVoteRank, and NCVoteRank, consider only the 1-hop (direct) neighbors of a node for the voting process to decide the influential node in each iteration. The proposed improved WVoteRank algorithm considers 1-hop neighbors and 2-hop neighbors for the voting process, which adequately covers the region, leading to an overall better selection of spreader nodes. The VoteRank and NCVoteRank methods are applicable only for undirected whereas WVoteRank method is applicable for weighted and unweighted network both. Our method applies to both weighted and unweighted networks and is an improvement over VoteRank, WVoteRank, and NCVoteRank algorithms. The proposed algorithm can be utilized to find influential nodes on all kinds of complex networks, such as transportation networks, collaboration networks, power networks, and social networks. Many considered datasets obey power-law and preferential attachment distributions. In the results section, we present the performance of our method on 10 network datasets of different sizes and applications. Of these, eight datasets are real-life networks, and the remaining two are synthetic networks.

Computational Complexity of the Improved WVoteRank Algorithm We assume that the number of nodes in the network is n and that the total number of edges in the network is m. The analysis can be performed as follows: We need to initialize the voting ability of each node $v \in V$, which requires $O(n)$ time. The voting score of a vertex is obtained by adding the voting abilities of all adjacent nodes up to two hops. This requires $\sum_{v_i \in V} d_i + \sum_{v_j \in V} d_j = O(m + n)$ time, where, for a specific node (u) under consideration, the first expression considers the direct neighbors of u and the second holds for the adjacent nodes up to two hops from node u, and d_i denotes the degree of node v_i. Consequently, the running time of this phase is $O(n + m)$. To select the spreader node with the maximum voting score in each requires $O(n)$ time. We need to select c (where c is a constant number) spreaders. Therefore, the selection of all spreaders can be done in $O(cn)$ time. To update the voting ability of all adjacent nodes up to two levels from the chosen influential nodes requires $O(c\langle k^2 \rangle) = O(c * m^2/n^2)$ time, where $\langle k \rangle$ is the average degree of the vertices in the system. Therefore, the running time of the proposed method is $O(n + m + cn + c * m^2/n^2)$, where c is a constant and that is $\ll n$. Normally, most real-life networks are sparse; we can consider $O(m) = O(n)$. Therefore, the time complexity of the proposed approach is $O(n)$.

The time complexity of the WVoteRank method is also $O(n)$. Hence, in terms of computational complexity, the proposed algorithm and WVoteRank are linear in terms of input size.

4 Information spreading model and performance metrics

We next describe the information spreading model and performance metrics used.

4.1 Spreading model

In this study, we used the most common stochastic epidemic-based information propagation model, namely, the susceptible infected recovered (SIR) model [31, 32], to examine the performance of various methods of influence maximization. This model divides the entire population into three parts: susceptible (S), infected (I), and recovered

(R). The nodes that are yet to receive information from their neighbors are called susceptible. Infected nodes refer to nodes that propagate the information, and recovered nodes denote the nodes that have already circulated the information and further onward do not propagate. The SIR model takes three inputs: initial spreaders, infection rate (β), and recovery rate (γ). In this model, all nodes are initially liable to get infected, except for a few nodes that are in the infected state. After every step, the infected nodes infect their susceptible neighbors with probability β. They then enter the recovered stage with probability γ. Once the recovered stage is reached, they are immunized and cannot be infected again. We run the model several times and the results are averaged over the total number of iterations to avoid the randomness involved in the model.

In the SIR model, the infection rate (β) refers to the rate at which an infected node can infect its susceptible neighbors. Suppose $\beta = 0.10$; then an infected node can randomly infect 10% of its susceptible neighbors during information diffusion. Similarly, the recovery rate (γ) is the rate at which infected nodes are recovered. Similarly, $\gamma = 0.5$ implies that 50% of the randomly infected nodes at time t transition to the recovered state at time $t + 1$, and $\gamma = 1$ implies that all the infected nodes at time t enter the recovered state at time $t + 1$ during the SIR model simulation.

4.2 Performance metrics

To judge the performance of the proposed method along with other methods for selecting the influential nodes, we consider the following metrics:

(i) Infection scale ($F(t)$): During the information diffusion process under the SIR model, the numbers of infected and recovered nodes vary with time in the system. At any instant t, the infection scale ($F(t)$) is the sum of the number of vertices in the network that are infected at time t and the total number of recovered nodes until time t divided by the total number of nodes in the network. The infected scale is the most practical measure of the effectiveness of the spreader selection algorithm as it provides us with an idea of the reach of information in the network with respect to time, originating from the chosen seed nodes. The following equation is used to compute $F(t)$:

$$F(t) = \frac{n_{I(t)} + n_{R(t)}}{n} \tag{14}$$

where $n_{I(t)}$, $n_{R(t)}$, and n represent the number of infected nodes at time t, total number of recovered nodes till time t and total number of nodes, respectively.

Note that the infection rate and infection scale are different. The infection rate refers to the rate at which an infected node can infect its susceptible neighbors. In contrast, the infection scale refers to the sum of the fraction of the numbers of infected and recovered nodes in the system at any timestamp t during the information diffusion process.

(ii) Final infected scale ($F(t_c)$): The final infected scale is an important performance parameter that corresponds to the extent to which information has finally spread in the network. It is the fraction of the number of vertices in the system that is infected and then recovered in the next time during the simulation of the SIR model. Suppose that the SIR simulation finishes at time t_c. Then $F(t_c)$ enumerates all those nodes that became infected and then recovered in the following rounds until time t_c during the process of information dissemination. The following equation is used to calculate the final infected scale, $F(t_c)$:

$$F(t_c) = \frac{n_{R(t)}}{n} \tag{15}$$

where $n_{R(t)}$ and n denote the number of recovered nodes and the total number of nodes respectively.

(iii) Average distance between spreaders (Ls): A diverse spread of influential nodes across various parts of the network is also a desirable parameter to ensure influence maximization. To achieve better coverage of the information, ideally, the distance between selected spreaders should be as great as possible. The diversity in selected spreaders can be determined by calculating the length of the shortest path between spreader nodes v_i and v_j. If we assume that the distance between nodes v_i and v_j is D_{ij} and that the number of spreader nodes selected is c, then the average distance (Ls) between spreaders can be calculated using the following equation:

$$Ls = \frac{\sum D_{ij}}{\frac{c(c-1)}{2}} \tag{16}$$

5 Datasets and a toy network

To examine the proposed approach, along with other existing methods, we used complex network datasets of different natures and capacities. The selection of datasets includes various graph characteristics exhibited by the networks, which can fairly evaluate the algorithms performance. Some of the datasets considered are listed in SNAP [33], which is a social network dataset repository. Most of the datasets used in this study follow power-law degree distributions, where most of the nodes have a very low degree and few nodes have a high degree. A brief description of the datasets is presented in Table 1.

Table 1 Dataset description

Dataset	#nodes	#edges
Powergrid	4941	6594
Facebook like- Social Network	1899	20297
US top 500 Airport Network	500	28237
Bitcoin+11	5881	35592
Euroroad	1174	1417
Ca-GrQc	5242	14496
as-733	6474	13895
Bible	1773	16401
BA	2000	9974
PCG	2000	9963

1. Powergrid [34]: This is a weighted network containing details of the power grid of the western states of the USA.
2. Facebook-like Social Network [35]: This is a weighted dataset compiled from a student's online community at the University of California, Irvine. The dataset comprises users who sent or received at least one message.
3. US Top-500 Airport Network [36]: This is a weighted network of the 500 busiest airports in the USA. The weights refer to the number of seats available on the proposed flights.
4. Bitcoin+11[37]: This is a trust network of persons who trade using Bitcoin on a platform called Bitcoin OTC. The participants of this network rate other members in a range of −10 (total distrust) to +10 (total trust).
5. Euroroad [38]: This is an international E-road network of roads primarily located in Europe. The network is undirected, and nodes represent cities, and a link between two cities represents a road.
6. Ca-GrQc [39]: This is an undirected collaboration network between Arxiv General Relativity and Quantum Cosmolog. The dataset includes an undirected edge between u and v if researcher u has co-authored a paper with researcher v.
7. Bible [40]: This is an undirected network containing nouns of the King James Version of the Bible and information about their co-occurrences. A node represents one such noun and an edge corresponds to two nouns that appear together in the same version of the Bible.
8. as-733 [41]: This is a system of routers and the Internet; it is designed into subgraphs called autonomous systems. It is a type of communication network dataset.
9. BA [42]: This is a random graph generated using the Barabási-Albert (BA) preferential attachment model. According to this model, a graph of n nodes is produced by adding new nodes, each with m edges that are preferentially attached to existing nodes with a high degree. Here, we use the value of n as 2000 and that of m as 5, and finally the generated graph contains 9974 edges. In this random graph, the weight of each added edge is a random number between 1 and 20.
10. PCG [43]: Holme and Kim introduced the power-law cluster (PCG) method for generating random graphs with a power-law degree distribution and approximate average clustering. The method takes three parameters (n, m, p), where n is the number of nodes, m is the number of random edges to be added for each new node, and p is the probability of adding a triangle after adding a random edge. The weight of each added edge is a random number between 1 and 20. We use the value of n as 2000, m as 5, and p as 0.5, and generated a synthetic graph containing 9963 edges.

We used a toy network consisting of 13 nodes depicted in Fig. 2 to understand the functioning of our algorithm in a lucid manner. The toy network is constructed in such a way that it enables us to compare the score of each node and the selection of spreader nodes by each method. Simulating the proposed method on a toy network selects node B as a source spreader in the first iteration; however, weighted h-index, WVoteRank selects node A and weighted degree selects node H as the source node. The selection of node B as a source spreader by the proposed method reveals its capability to elect a distinct node as a seed node.

Table 2 lists the scores of each node using different methods such as weighted degree, weighted closeness, weighted betweenness (W-Betw.), weighted h-index, WVoteRank

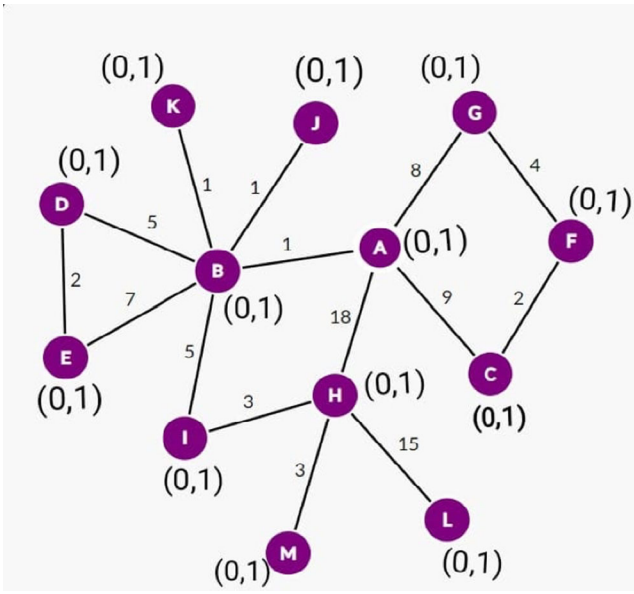


Fig. 2 Toy network after the first iteration. The vote score obtained and voting ability of each node using the proposed approach are depicted in the diagram

Table 2 The score of various nodes of the toy network using distinct approaches

Node	W-Degree	W-Closeness	W-Betw.	W- *h*index	WVR	I-WVR
A	20.784	0.122	0.409	**66**	12	21
B	14.142	0.129	0.803	66	10	**22**
C	4.490	0.071	0.151	36	4	10
D	2.857	0.086	0.083	42	3	8
E	3.674	0.075	0.0	40	4	8
F	2.449	0.064	0.015	16	3	7
G	4.898	0.075	0.0	36	4	10
H	22.516	0.086	0.318	49	12	20
I	3.265	0.101	0.409	50	4	11
J	0.288	0.115	0.0	30	1	4
K	0.288	0.115	0.0	30	1	4
L	4.330	0.039	0.0	20	3	7
M	0.866	0.069	0.0	20	1	6

Bold entries in each row signify the highest value obtained by a particular method

(WVR), and improved WVoteRank (I-WVR), which equips us to know the best node under each process.

6 Results and analysis

We performed our investigation using all the performance measures presented in Section 4.2. We compared the performance of the proposed approach against that of popular centrality measures such as weighted degree [12], weighted closeness calculated using (3), weighted betweenness computed using (4), weighted h-index [17], and weighted VoteRank [29]. The investigation results were examined on eight real-life and two synthetic weighted and unweighted datasets of diverse sizes and properties. The datasets used are listed in Table 1: Powergrid is a power network, Facebook is a social network, US Airport is a transportation network, Bitcoin is a trust network, Ca-GrQc is a collaboration network, and as-733 is a kind of communication network. As can be seen, the collection of used real-life datasets belonged to different types of complex networks. Most of the datasets used in the paper, such as as-733, Ca-GrQc, Bible, BA, and PCG, follow a power-law distribution. We utilized the stochastic SIR model to calculate the infection scale ($F(t)$) and final infected scale ($F(t_c)$) as a function of time (t). We executed the SIR simulation 100 times, and the results were averaged over the total number of times for each performance measure. For consistency and to maintain coherence in the evaluation of all datasets, we considered the value of infection rate (β) equal to 0.01, implying that, once a node gets infected at any timestamp, it can randomly infect 1% of its neighbors in the next timestamp. During the execution of the SIR model, we take the value of the recovery rate (γ) as 1. This implies that, if a node is infected at time t, then it is recovered at the next time, $t + 1$; furthermore, it has no role in information propagation.

6.1 Infection Scale ($F(t)$) vs. Time

Figure 3 presents the infected scale, $F(t)$, with respect to the time for the different network datasets, as depicted in Table 1 under a given infection rate (β) and number of source spreader nodes. The value of $F(t)$ is calculated by applying (14). The infected scale is an indicator of how promptly a message spreads in the network starting from the source node. To maintain consistency in the assessment, we used the value of infection rate (β) as 0.01 and the number of source influential nodes as 100 for all the datasets. Strategically chosen source nodes can achieve high infection levels promptly. From Fig. 4, it is evident that the proposed improved WVoteRank approach outperforms all other models such as weighted degree, weighted closeness, weighted betweenness, weighted h-index, and weighted VoteRank on both weighted and unweighted networks.

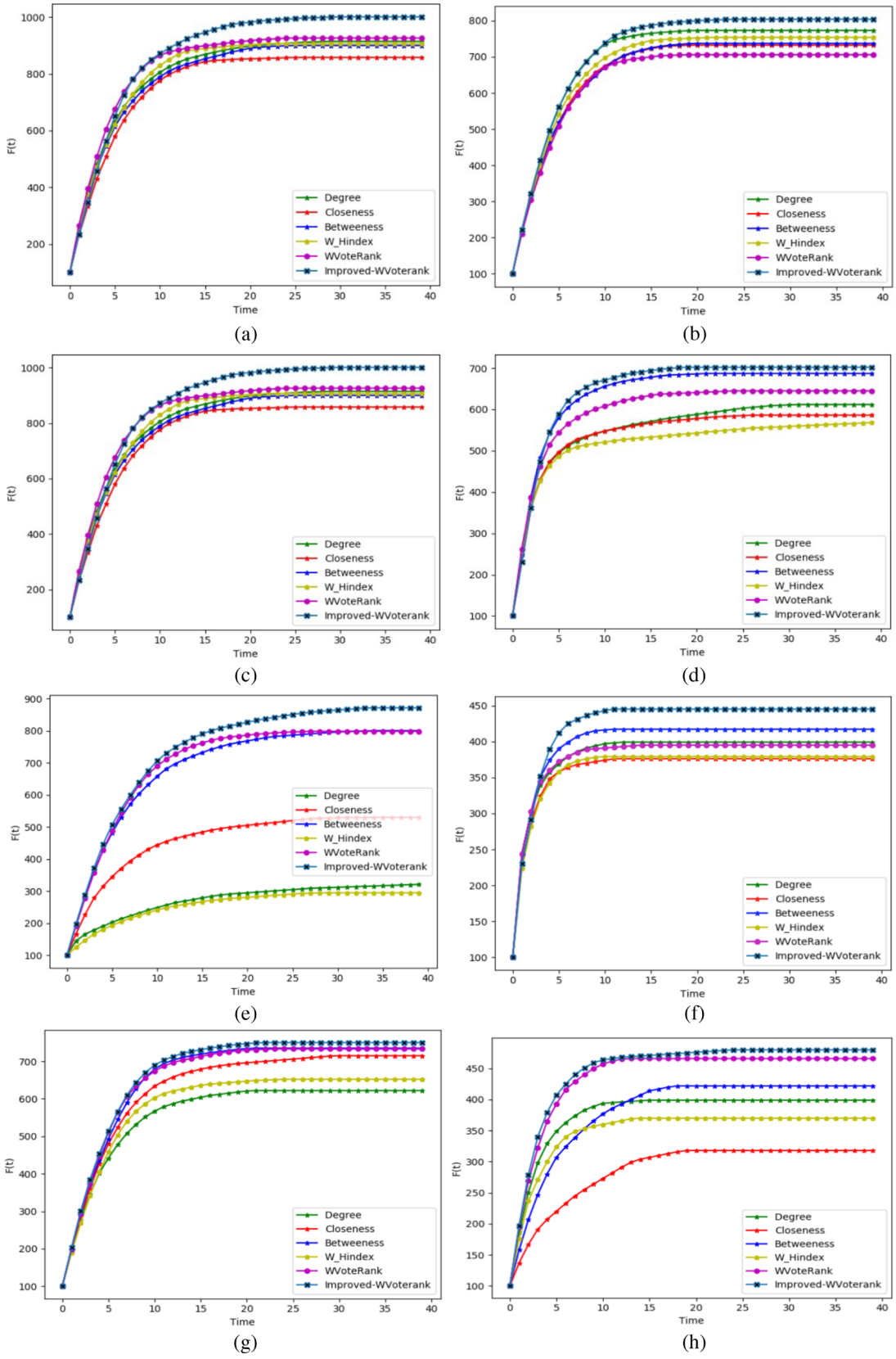


Fig. 3 (a)-(h): Infected scale (F_t) Vs. time of different real-life datasets with the same value of β as0.01 and 100 initial spreaders for all datasets

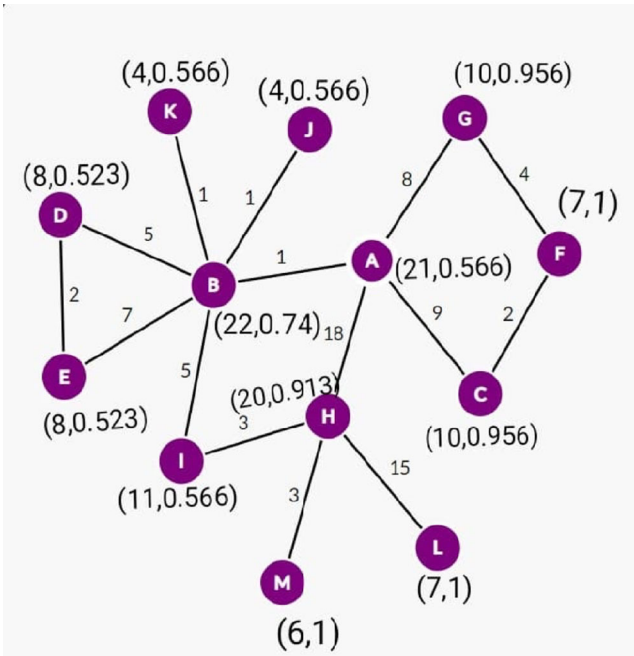


Fig. 4 Toy network after the second iteration

6.2 Final infected scale ($F(t_c)$) vs. time

In Fig. 5, we illustrate the value of $F(t_c)$ with respect to time for all the datasets using a given infection rate and the number of seed nodes. We computed the value of $F(t_c)$ using (15). The value of the final infected scale indicates the portion of the system that has received the information originating from the seed nodes. Intuitively, the higher the value of the final infected scale, the better the influence maximization. Similarly to our treatment of the infection scale ($F(t_c)$) vs. time, we considered the value of infection rate (β) as 0.01 and the number of initial spreaders as 100 for all the datasets. From the final infected scale($F(t_c)$) performance matrix, it is quite clear from Fig. 5 that our proposed method consistently performs better than the previously proposed weighted VoteRank, weighted h-Index, weighted betweenness centrality, weighted degree centrality, and weighted closeness centrality methods. Figure 6 depicts the results obtained on two popular random graphs, namely, BA and PCG, where the proposed algorithm also yields comparably better results. Therefore, we can anticipate improved weighted VoteRank to exhibit its domination in larger sized real-life and synthetic graphs.

6.3 Average shortest path length between spreaders (L_s)

The extent to which the information can finally reach the network, starting from the source influential nodes, also depends on their relative positions. When the chosen spreaders are elected from the distinct positions within the network, this can lead to maximum coverage. Table 3 lists the average shortest path length (Ls) between the elected spreaders by employing (16) for different datasets using all the methods. Similar to our treatment of the infection scale ($F(t)$) vs. time and final infected scale ($F(t_c)$) vs. time, we used the value of infection rate (β) as 0.01 for the measure of L_s. From Table 3, it is clear that we are able to reach distant nodes and maximize influence spreading through the improved weighted VoteRank method. Our method produces the best result for the considered real-life datasets, i.e., Bitcoin+11, Facebook-like Social Network, Powergrid, US Top-500 Airport network, Ca-GrQc, as-733, Bible, and Euroroad. The proposed algorithm also outperforms other algorithms on two synthetic datasets, namely, BA and PCG.

Based on the obtained results for different measures such as $F(t)$ vs. time (3), $F(t_c)$ vs. time (Figs. 5 and 6), and Ls between spreaders (Table 3), it is evident that the proposed method identifies influential nodes from different places, leading to better influence maximization. Owing to the proper topological location in the network, influential nodes perform an essential role in the fast propagation of information in the system.

7 Conclusion

Influence maximization in a complex network is achieved by judiciously selecting influential nodes. Such nodes are capable of spreading information to the maximum number of nodes in the network. Finding influential nodes in a weighted network is more challenging than in an unweighted network. In a weighted network, the edges between nodes bear weights depicting their strength or capacity. In this study, we proposed a novel method called improved WVoteRank to identify and rank the nodes based on their spreading capabilities in weighted networks. Furthermore, our method is applicable to unweighted networks as well. The improved method identifies influential nodes from the distinct positions in

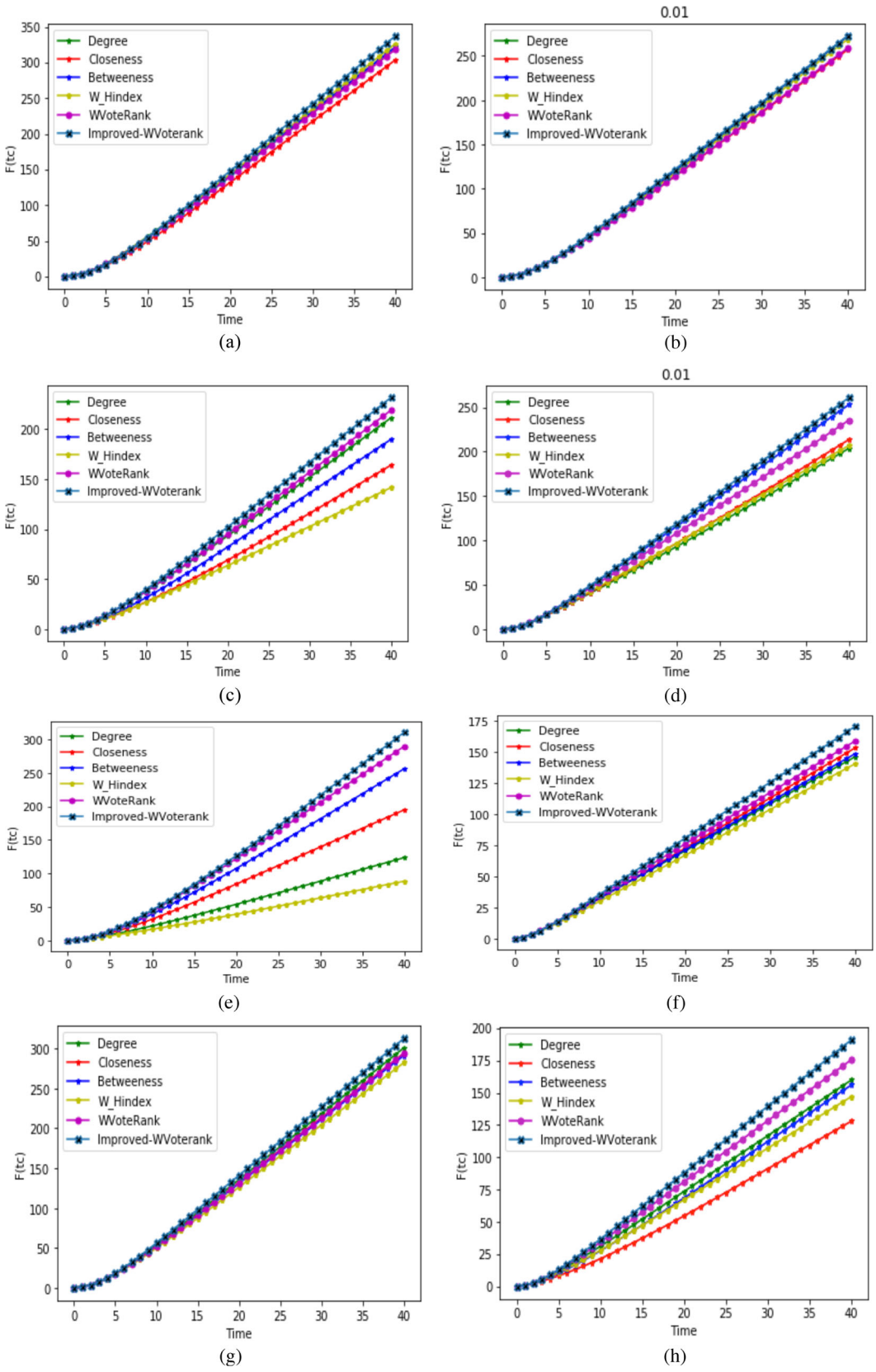


Fig. 5 Final Infected scale with respect time for real-life datasets with same infection rate (β) = 0.01 and initial number of spreaders as 100

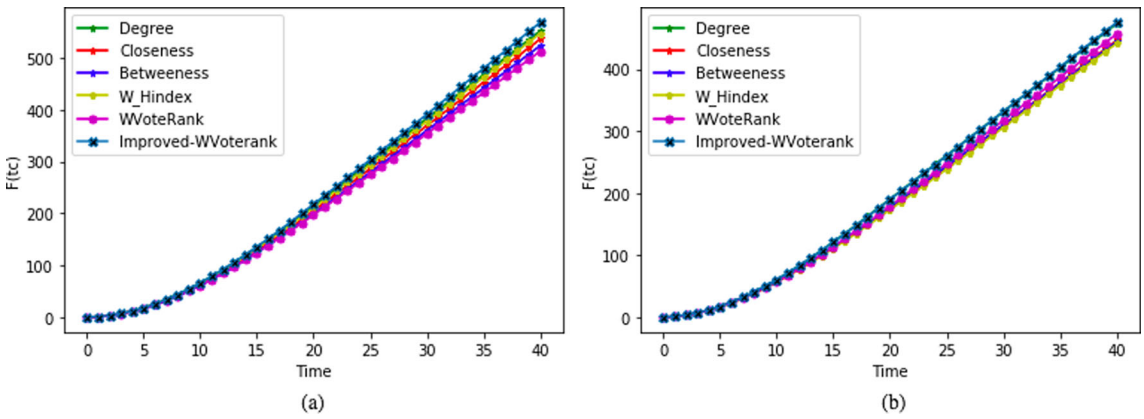


Fig. 6 Final Infected scale with respect time for two synthetic datasets, namely Barabasi-Albert (BA) and Power-law cluster graph (PCG) with infection rate (β) = 0.01 and an initial number of spreaders as 100

the system and has a linear running time in terms of the input dataset size. In the case of the WVoteRank method, only 1-hop neighbor nodes contribute to the voting process. However, based on the extended neighborhood concept, we argue that nodes other than direct neighbors can also play an important role in the selection of top spreaders. In each round of voting, the node with the highest vote score is selected as an influential node. We updated the voting ability of its neighbors up to two hops. Furthermore, we examined the performance of various methods on eight real-world datasets. The final infected scale ($F(t_c)$) and average path length between spreaders (L_s) reported by our approach exhibited notable improvement over values obtained by other methods available in the literature. We attained enhanced performance by the distinct selection of source nodes. In addition, increasing the number of source nodes was consistently maintained the performance of the proposed algorithm. As future work, we intend to extend our proposed algorithm to identify influential spreaders in multilayer weighted complex networks. In complex weighted networks like social networks, people or nodes may belong to more than one social network. Therefore, it is worth examining these networks by acknowledging the overlapping properties of nodes to map real-life situations. We will also investigate the IWVR method to find opinion leaders in social networks.

Table 3 Average Shortest Path Length (Ls) values between elected spreaders with infection rate(β) =0.01 and initial spreaders fraction=0.05 where WD, WC, WB, WH, WVR, $IWVR$ represents weighted-degree, weighted-closeness, weighted-betweenness, weighted h-index, WVoteRank and improved-WVoteRank method respectively

Dataset	$Ls(WD)$	$Ls(WC)$	$Ls(WB)$	$Ls(WH)$	$Ls(WVR)$	$Ls(IWVR)$
POWERGRID	334.630	89.554	130.489	216.695	290.347	336.760
BITCOIN	25.566	20.275	22.583	24.465	26.460	43.504
FACEBOOK	33.277	30.583	33.194	31.527	32.027	40.916
US AIRPORT	60.285	37.714	45.0	51.333	56.0	252.714
EUROROAD	7.0	2.0	3.2	5.9	8.8	12.2
CA-GRQC	29.846	47.446	53.415	20.0	64.061	63.692
AS-733	32.379	29.395	32.862	29.112	34.153	57.822
BIBLE	17.892	18.321	18.142	17.642	18.464	30.321
BA	69.77	65.16	67.87	67.44	70.39	84.63
PCG	68.31	54.30	65.16	61.21	68.69	82.91

References

1. Newman MEJ (2010) Networks: an introduction. Oxford University Press, New York
2. Wang XF, Li X, Chen GR (2012) Network Science: An introduction, Higher Education Press, Beijing, pp 10–15
3. Newman MEJ (2004) Analysis of weighted networks. Phys Rev E 70(5):056131
4. Kandhway K, Kuri J (2016) Using node centrality and optimal control to maximize information diffusion in social networks. IEEE Trans Syst Man Cybern Syst 47(7):1099–1110
5. Li Y, Fan J, Wang Y, Tan KL (2018) Influence maximization on social graphs: a survey. IEEE Trans Knowl Data Eng 30(10):1852–1872
6. Ferguson R (2008) Word of mouth and viral marketing: taking the temperature of the hottest trends in marketing. J Consum Market 25(3):179–182
7. Sun Y, Liu C, Zhang CX, Zhang ZK (2014) Epidemic spreading on weighted complex networks. Phys Lett A 378(7):635–640
8. Wang W, Tang M, Eugene Stanley H, Braunstein LA (2017) Unification of theoretical approaches for epidemic spreading on complex networks. Rep Progr Phys 80(3):1–16
9. Pastor -Satorras R, Castellano C, Van Mieghem P, Vespignani A (2015) Epidemic processes in complex networks. Rev Modern Phys 87(3):925–979
10. Freeman LC (1978) Centrality in social networks conceptual clarification. Soc Netw 1(3):215–239
11. Lu LY, Chen DB, Ren XL, Zhang QM, Zhang YC, Zhou T (2016) Vital nodes identification in complex networks. Phys Rep 650:1–63
12. Opsahl T, Colizza V, Panzarasa P, Ramasco JJ (2008) Prominence and control: the weighted rich-club effect. Phys Rev Lett 101(16):168702
13. Freeman LC (1977) A set of measures of centrality based on betweenness. Sociometry 40(1):35–41
14. Sabidussi G (1966) The centrality index of a graph. Psychometrika 31(4):581–603
15. Hirsch JE (2005) An index to quantify an individual's scientific research output. Proc Natl Acad Sc 102(46):16569–16572
16. Lu L, Zhou T, Zhang QM, Stanley HE (2016) The H-index of a network node and its relation to degree and coreness. Nat Commun 7:10168
17. Yu S, Gao L, Wang YF, Gao G, Zhou C, Gao ZY (2017) Weighted H-index for identifying influential spreaders. arXiv:1710.05272
18. Kitsak M, Gallos LK, Havlin S, Liljeros F, Muchnik L, Stanley HE, Makse A (2010) Identification of influential spreaders in complex networks. Nat Phys 6(11):888–893
19. Zeng A, Zhang CJ (2013) Ranking spreaders by decomposing complex networks. Phys Lett A 377(14):1031–1035
20. Bae J, Kim S (2014) Identifying and ranking influential spreaders in complex networks by neighborhood coreness. Physica A: Stat Mech Appl 395:549–559
21. Garas A, Schweitzer F, Havlin S (2012) A k-shell decomposition method for weighted networks. J Phys 14(8):083030
22. Yang F, Li X, Xu Y, Liu X, Wang J, Zhang Y, Zhang R, Yao Y (2018) Ranking the spreading influence of nodes in complex networks: an extended weighted degree centrality based on a remaining minimum degree decomposition. Phys Lett A 382(34):2361–2371
23. Wang F, Jiang W, Li X, Wang G (2018) Maximizing positive influence spread in online social networks via fluid dynamics. Futur Gener Comput Syst 86:1491–1502
24. Berahmand K, Bouyer A, Samadi N (2019) A new local and multidimensional ranking measure to detect spreaders in social networks. Computing 101(11):1711–33
25. De Souza RC, Figueiredo DR, Rocha ADA, Ziviani A (2020) Efficient network seeding under variable node cost ans limited budget for social networks. Inf Sci 514:369–384
26. Newman MEJ (2004) Fast algorithm for detecting community structure in networks. Phys Rev E 69(6):066133
27. Huang H, Shen H, Meng Z, Chang H, He H (2019) Community-based influence maximization for viral marketing. Appl Intell 49(6):2137–50
28. Zhang J. X., Chen DB, Dong Q (2016) Identifying a set of influential spreaders in complex networks. Sci Rep 6(27823):p1–p9
29. Sun H L, Chen DB, He JL, Cheng E (2019) A voting approach to uncover multiple influential spreaders on weighted networks. Phys A: Stat Mech Appl 519:303–312
30. Kumar S, Panda BS (2020) Identifying influential nodes in Social Networks: Neighborhood Coreness based voting approach. Physica A: Stat Mech Appl:124215
31. Pastor-Satorras R, Vespignani A (2001) Epidemic dynamics and endemic states in complex networks. Phys Rev E 63(6):066117
32. Hethcote HW (2000) The mathematics of infectious diseases. SIAM Rev 42:599–653
33. Leskovec J, Krevl A (2014) SNAP Datasets: Stanford Large Network Dataset Collection
34. Watts DJ, Strogatz SH (1998) Collective dynamics of small-world networks. Nature 393:440–442
35. Opsahl T, Panzarasa P (2009)) Clustering in weighted networks. Soc Netw 31(2):155–163. https://doi.org/10.1016/j.socnet.2009.02.002
36. Colizza V, Pastor-Satorras R, Vespignani A (2007) Reaction-diffusion processes and metapopulation models in heterogeneous networks. Nat Phys 3:276–282
37. Kumar S, Spezzano F, Subrahmanian VS, Faloutsos C (2016) Edge Weight Prediction in Weighted Signed Networks. IEEE International Conference on Data Mining (ICDM)
38. ?ubelj L, Bajec M (2011) Robust network community detection using balanced propagation. Eur Phys J B 81(3):353–362
39. Leskovec J, Kleinberg J, Faloutsos C (2007) Graph evolution: Densification and shrinking diameters. ACM Trans Knowl Discov Data (ACM TKDD) 1(1)
40. Bible network dataset (2017) KONECT
41. Leskovec J, Kleinberg J, Faloutsos C (2005) Graphs over time: Densification Laws, Shrinking Diameters and Possible Explanations. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)
42. Barabási AL, Albert R (1999) Emergence of scaling in random networks. Science 286(5439):509–12
43. Holme P, Kim BJ (2002) Growing scale-free networks with tunable clustering. Phys Rev E 65(2):026107

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Sanjay Kumar is currently an Assistant Professor in the Department of Computer Science and Engineering, Delhi Technological University, New Delhi, India. He is pursuing a Ph.D. degree in Computer Science and Application group of the Department of Mathematics, Indian Institute of Technology (IIT) Delhi. He has completed M. Tech in Computer Application from Indian Institute of Technology (IIT) Delhi, India. His research interests include Social Network Analysis, Graph Algorithm, Machine Learning, and Design and Analysis of Algorithms.



Ankit Panda did his Bachelors of Computer Science and Engineering from Bennett University, India. He is currently pursuing Masters of Machine Learning from Department of ECMS (Engineering, computer and Mathematical Sciences), the University of Adelaide, Australia. He has a keen interest in Machine learning and has worked in financial data in the University of Western Switzerland, HEIG-VD as an Exchange student as part of his B.Tech.

Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021)
IEEE Xplore Part Number: CFP21K74-ART; ISBN: 978-0-7381-1327-2

Image Detection in Noisy Images

Kushagra Yadav
Computer Science and Engineering
Delhi Technological University
New Delhi, India
kushagra_bt2k17@dtu.ac.in

Dakshit Mohan
Computer Science and Engineering
Delhi Technological University
New Delhi, India
dmmaheshwari.88@gmail.com

Anil Singh Parihar
Computer Science and Engineering
Delhi Technological University
New Delhi, India
parihar.anil@gmail.com

2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) | 978-1-6654-1272-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICCS51141.2021.9432243

***Abstract*—The problem of image noise is prevalent in various visual recognition tasks. This problem arises due to multiple factors like insufficient illumination, low-quality camera sensors, background-clutters, etc. Low-quality images incur degraded performance for visual recognition tasks like classification, object detection, because of the noise present in the image. Thus, mostly the latest image recognition benchmarks and methods, such as Pascal Visual Object Classes Challenges and Microsoft Common Objects in Context Challenges, rely primarily on clean and high-definition images. Furthermore, image noise is a typical problem in real-time applications like surveillance, autonomous vehicles, etc. To tackle this bothersome problem of underperformance of visual recognition tasks, particularly object detection, due to inclusion of noise in real-time scenarios, a technique has been proposed to reduce the image noise's impact on the object detection task and improve its performance altogether. A two-phase pipeline is introduced, in which the first phase consists of removing noise from the low-quality image, and the second stage includes accomplishing the object detection process by employing the standard Single-shot multibox object detector. The proposed approach is evaluated on the Pascal Visual Object Classes benchmark. The detection results and image denoising results exhibit the efficiency of the proposed technique for reliable object detection with varying image noise levels.**

***Keywords*—Image denoising, Object detection, Deep learning, Residual learning**

I. INTRODUCTION

The task of object detection requires considerable efforts—the reason being the significant variations occurring in the images related to view-points, background clutter. Furthermore, illumination changes, unfavorable weather conditions can lead to reduced detector performance [1]. High-quality images depict a significant job in scenarios where the object-detection task has great importance. For instance, autonomous driving cars require object detection algorithms to operate accurately and consistently. Low-quality images may cause a hindrance in their smooth operation.

The vision community has done extensive research related to this subject. Girshick *et al*. [2] were the first to introduce the use of CNNs in object detection. They proposed the approach of employing region proposal network before computing features and classification network. In 2015, Redmon *et al*. [3] introduced the idea of 'only looking once' at the image while processing called YOLO. They repurpose the object detection process as a problem of regression as an alternative for any classification problem. In [4], following the strategy as [3], Liu *et al*. proposed a multibox detector to process the object detection task in a single pass of the network.

Dealing with noise can sometimes be tedious. In [5], Boyat and Joshi explain the fundamental noise models in digital image processing. Currently, Gaussian noise model is the most used image noise model. Here, it is more concerned about the effects of this Gaussian noise (AWGN) on object detection. In the past, researchers proposed various image denoising models with image priors, such as BM3D [6], NCSR [7], and NSS [8]. However, these methods had limitations like high computational cost and non-convex nature. To address these limitations, several CNN-based denoising techniques were introduced in recent years [9]–[12]. Initially, [9] was the first work to introduce the idea of using CNN in image denoising. In [11], Liu *et al*. implemented a wavelet transform on top of U-net architecture to denoise images. Furthermore, Zhang *et al*. introduced the concept of residual learning in image denoising [12].

One can think of ways to minimize the impact of digital noise in image recognition tasks. For instance, one idea is to denoise the image before applying object detection. Another alternative can be to re-train the object detection network on the noisy images, learn the features from that noisy image, and produce improved results on the re-trained network. This research work emphasizes on the above approach to first denoise the images and then employ an object detector on those preprocessed images in this paper.

Our first phase includes the image denoising task via residual learning, which is fulfilled by using the Residual Dense Network (RDN) [12] method. This method exploits the hierarchical features with different receptive fields from all the network's convolutional layers to restore the original high-quality picture. So, final output of this RDN network is a denoised image. Following the preprocessing, the noise-free image goes into the Single Shot Multibox Detector (SSD) network [4] for the detection process. The detector then predicts the categories of the objects in the image and localizes them as well, that is, predicts their bounding boxes. The obtained detection results are reported on the Pascal VOC2007 dataset [13], and the results are also demonstrated on the images with varying levels of noise.

The rest of the paper is organized as: In section2, there is the review of the related work on the concerned topics of image denoising and object detection. In Section 3, we propose

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:32:36 UTC from IEEE Xplore. Restrictions apply.

our methodology for improving the performance of object detector in case of images imbued with noise. We demonstrate our model's performance on Pascal VOC dataset. We then compare these denoising results with other denoising methods in Section 4. Lastly, in section 5, we provide the conclusion of our paper.

II. Related Work

Object Detection: The objective of accurate visual recognition has been under investigation for a long time. During the last decade, researchers have made considerable progress in using SIFT [14] and HOG [15] methods for various visual recognition tasks. Both methods are described as block-wise orientation histograms, a rough representation of cells present in the primary cortical region of a primate's optical path. Other classical works such as [16], and [17] considered object detection as a classification problem solved by a sliding window approach.

However, by introducing deep convolutional neural networks (CNNs) for image recognition tasks [18] in 2012, Krizhevsky *et al*. brought a revolution in the computer vision community. They demonstrated a noteworthy growth on the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [19], [20] having the classification accuracy as the metric. Since then, several CNN based approaches [2]–[4], [21], [22] have been authored for detection tasks. Among these, SSD [4] and YOLO [3] received greater attention from the community due to their fast speed of detection and less prediction time. Prior works such as R-CNN [2], Fast R-CNN [22], and Faster R-CNN [21] involve a step where they produce object proposals with a network called Region Proposal Network (RPN) and then classify each region independently. Therefore, these complex detection pipelines are slow and computationally expensive. So, Redmon *et al*. presented YOLO [3] and Liu *et al*. proposed SSD [4] to deal with these limitations. Both approaches require only a single forward pass for the detection process and, hence, are single-shot detection models. The main difference between both methods is in their architecture. YOLO network consists of 24 convolutional layers, followed by two densely connected layers, which finally predict the categories and bounding boxes. On the other hand, the SSD model appends extra feature layers to the base network, resulting in bounding boxes prediction with different scale and aspect ratio. SSD performs better than YOLO in terms of accuracy or exactness, whereas YOLO is better than SSD in terms of speed. In this paper, we adopt SSD as the object detector because of its modest accuracy-speed trade-off.

Image Denoising: The denoising problem is of great concern for last two decades. In the early 90s, researchers proposed techniques such as wavelet coring [23], total variation denoising [24], and anisotropic diffusion [25]. These methods employ custom manually crafted algorithms and thus, retrieve a clear signal from non-clear input. In the next decade, techniques driven by self-similarities, such as BM3D [6] and non-local means [26], became prominent in the vision community. BM3D uses a 3-D block-matching algorithm for noise reduction in images.

In 2009, Jain and Seung [27] introduced the application of CNN for denoising of images, resulting in a paradigm shift. Recently, many researchers are adopting multi-layer perceptron (MLP) networks to train them to calculate the conversion from noise patch to a clean pixel. Zhang *et al*. proposed DnCNN [9] in which the network determines the mappings from noisy image to the noise by implementing residual learning and batch normalization [28]. Zhang *et al*. in IRCNN [29] improved the technique by incorporating dilated filtering in the network and learning CNN denoiser prior to the restoration task. Recently, Liu *et al*. in [11] presented MWCNN for the tradeoff between receptive field size and computational complexity and employing wavelet transform to reduce the feature map size. In [12], Zhang *et al*. introduced RDN, the idea of learning the hierarchical features generated by a neural network and utilizing them for problems like image denoising, image super resolution, etc. Here, we utilize RDN in image denoising phase as the network produces quality results with reduced network depth and high computational efficiency as compared to other models like MemNet [30].

III. Proposed Methodology

In this section, we discuss our approach to producing a real time and reliable object detection system.

A. Model Pipeline

We accomplished the designated object detection task by designing a two-step pipeline. In the first stage of this pipeline, by training a neural network, we reduce the noise present in the low-quality input image. In the second step, by training another model having the denoised image as the input, we detect the objects present in the image and localize them. We reveal you our proposed model pipeline's overview in figure 1.

Our objective is to clear the noise that is weakening the image's quality in the first step of the pipeline. We implemented an image denoising network introduced by Zhang *et al*.
[12] known as Residual Dense Network (RDN) to achieve our goal. The network reduces an image's noise by mapping the noised image to a high-quality through the concept of residual learning.

RDN consists of a set of residual dense blocks (RDBs) that have dense connections of convolutional layers. Each layer in an RDB is densely connected with its preceding layers. Furthermore, RDB enables uninterrupted links from previous residual dense block to every convolutional layer in the current block, thus producing continuous memory mechanism and local feature fusion. We know that a larger growth rate contributes to better performance but also hinders the training as the network becomes deeper. The local feature fusion of the RDB along with local residual learning overcomes this issue by adaptively preserving the information and utilizing all the hierarchical features from the original image. RDN also extracts the global hierarchical features from the input by

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:32:36 UTC from IEEE Xplore. Restrictions apply.

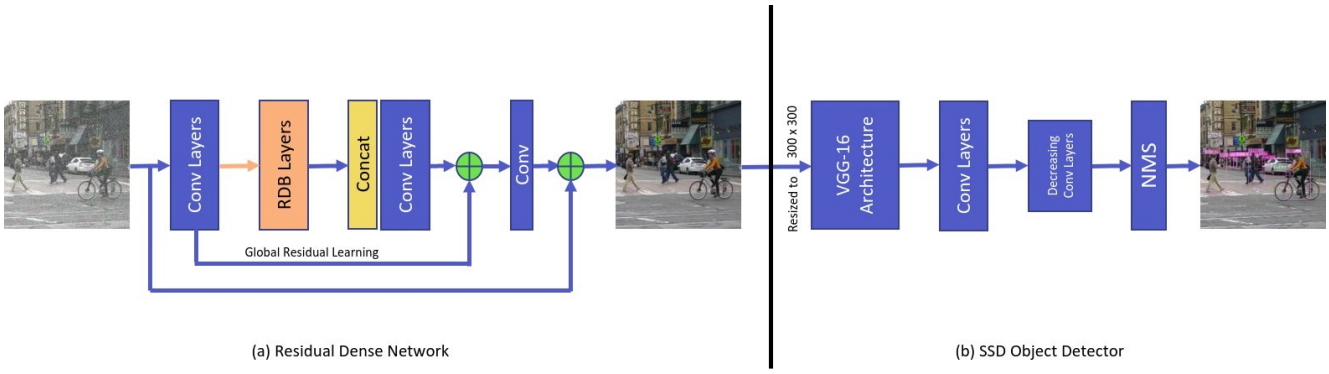


Fig. 1: **Overview of our detection pipeline**. We tackle the problem of reduced detection performance due to low-quality images, in two stages. In the first stage, the model, with the use of subsequent convolutional layers and via residual learning, removes the image noise. In the second stage, the network predicts the category scores and bounding boxes of objects in the image using the SSD object detector.

concatenating features from every RDB and utilizing global residual learning. This concatenation adaptively fuses the multi-level local dense features from the RDBs. The end product of the network is a high-quality denoised image.

We put forward the following arguments on selecting RDN [12] over other denoising networks like MemNet [30] and SRDenseNet [31]. The RDN doesn't make use of any pooling layer as they may remove pixel-level features which are important for image details. The concatenation of local features obtained from each RDB prove to be more effective than the transition layers between dense blocks in networks inspired from DenseNet. The use of these local features also stabilizes the training when the network becomes deep while using a larger growth rate. Furthermore, the use of global feature fusion, which utilize the hierarchical features, is not seen in other dense network implementations. Other networks tend to use L2 loss which is good for maximizing peak signal-to-noise ratio. However, L1 loss provides better convergence and helps RDN to achieve better performance. RDN doesn't utilize any batch normalization layer and works on the low-quality image with its original size, thus, reducing computational complexity and consuming fewer hardware resources.

The next step in our pipeline is the object detection task. For this purpose, we integrated a Single Shot Multi-box detector (SSD) in our pipeline. This object detection algorithm predicts object classes and their location of bounding box in an image utilizing only a single pass through the neural network. The detector approaches the object detection intricacy as a regression problem. Usually, the intricacy is approximated as a classification problem. It discretizes bounding boxes' output space in a collection of anchor boxes for varying aspect ratios and varying object scales. It also takes care of the problem of varying object sizes in the image. Since SSD condenses the whole enumerations into a sole network, it is highly efficient as compared to other detectors like the family of R-CNNs. We employed the SSD as the detector in our pipeline as it is trained easily and straightforward for integration with methods requiring object detection.

B. Network Architecture

The two modules that make up our model are trained separately as they together constitute a complex network. Training them individually and utilizing them in a single two-stage pipeline reduces the overall complexity of the pipeline. Residual Dense Network takes a noisy input image and learns hierarchical features in low-quality space for noise removal. Then the denoised outcome goes into the SSD network. It determines filters to regress bounding box offsets in the image.

Residual Dense Network: The denoising network mostly comprises of three parts: dense feature fusion (DFFs), residual dense blocks (RDBs), and shallow feature extraction (SFE). It uses two convolutional layers to obtain superficial image features. Additionally, second convolutional layer is used for global residual learning. The network contains various residual dense blocks, let the number be D. The RDB includes a set of many, say C, convolutional, and ReLU [32] layers. The individual layer in RDB is connected to its preceding layer to extract abundant local hierarchical features. Each RDB's output is interpreted as a local feature, and these features are conveyed to the next RDB and its constituent layers to form a contiguous memory mechanism. Finally, all the local features from RDBs are concatenated. These concatenated features are then forwarded into a convolutional layer to determine the global features from the network. The global features are then added to the shallow features extracted via residual learning to fuse the dense features. These fused dense features go in the final Conv layer to predict a high-quality image.

SSD Object Detector: SSD predicts a fixed-sized set of bounding boxes and scores for the objects present. SSD is built on the top of VGG-16 [33] as the base network but discards the fully connected layers, and various Conv layers are appended to it to predict detections. The reason for choosing VGG-16 as the base network is because it provides outstanding performance in image classification taks, and thus, is utilized

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:32:36 UTC from IEEE Xplore. Restrictions apply.

to provide the classification scores required in the end results. In the VGG-16 architecture, three blocks of two convolutional and a pooling layer, and two blocks of three convolutional and a pooling layer are present. The filter size in this architecture is always 3 x 3 with number of filters increasing from 64 to 512. After the VGG-16 architecture, a convolution of 3 x 3 x 1024 is employed followed by a convolution of 1 x 1 x 1024. Similarly, layers with 1 x 1 and 3 x 3 size of kernels are used with 512 and 256 as the numbers of filters. The size of the filters is reduced from 3 x 3 x 1024 all the way to 3 x 3 x 256. These Conv layers are gradually decreasing in size to allow predictions of detections at multiple scales. Individual feature layers at different stages of the SSD network have different convolutional models, such as the 7^{th} Conv layer is the product of a single convolution operation. Foreach feature layer, the 3 x 3 *small kernels* in the convolution step produce either a category score or a bounding box location offset to the default bounding box location. In contrast, the 8th Conv layer is constructed from two convolution operations. In the end, following the detections predicted by the added feature layers, the non-maximum suppression algorithm is employed to deliver the final outputs.

C. Training and Inference Details

Both the modules-RDN and SSD are trained individually and end-to-end. In the first stage of the pipeline, we train RDN on noisy images to learn the hierarchical features to map the images to a better, noise-free image. Through the combination of various convolutional layers, the network learns local as well as global features. A single training sample of our model comprises of determining the noise-free patch corresponding to a reference noised image patch. We used the L1 loss function on the predicted high-quality image (I_{pred}) and the reference (I_{gt}) clean RGB image to train the entire RDN network. Usually, researchers prefer L2 loss over L1 since L2 maximizes the peak signal-to-noise ratio (PSNR). On the other hand, L1 provides better convergence, as empirically discovered by Lim *et al*. [34].

$$L1 = \sum_{i=1}^{N} |I_{pred} - I_{gt}| \quad (1)$$

In the pipeline's second stage, we train the SSD object detector to produce object category scores and bounding box in the image. The network learns to regress bounding box location offset to default bounding box locations. We used the categorical cross entropy loss function over the C object categories as the confidence loss to train the entire SSD network. The categorical cross entropy loss is defined as follows-

$$CE = -log\frac{e^{s_p}}{\sum_j^c e^{s_j}} \quad (2)$$

where s_p is the predicted score for positive class and s_j is the predicted score and the predicted score respectively for each class j in C.

As for the RDN network training, we create noisy samples from corresponding clean RGB images by adding additive white gaussian noise (AWGN). We then use these high-quality images as labels for the training. Noisy samples were created for varying noise levels, so, for each noise level, we use the above method to create a training dataset, having the clean image as the label and the noisy image as the input. The RDN network is then trained on these datasets. Coming to the detector training, we train the SSD network on the combination of Pascal VOC2007 and Pascal VOC2012 [13], [35] training and validation sets.

Evaluation metric: We evaluate the predictions by the RDN network using the average peak PSNR values between the denoised image and the reference clean image, over the validation dataset to assess the quality of our trained module. High PSNR values are interpreted as good quality results when the denoised image is compared to the original image. On the other hand, the fundamental objective of the whole pipeline is object detection. Therefore, the entire network and the object detector simultaneously use mean average precision (mAP) as the evaluation metric to assess their performance.

IV. EXPERIMENTS

A. Dataset

We utilize the DIV2K dataset [36] for training the denoising part of our network. This dataset comprises of 1000 images – 800 for training, 100 for testing, and 100 for validation. We used the training set of DIV2K to train the RDN network. On the other hand, we use the combination of Pascal VOC2007 and 2012 training and validation sets, also called `trainval` set, to train the SSD step in our network.

To evaluate both the trained models of our network, we use the Pascal VOC2007 `test` set with 4952 images. For each high-quality validation image, we generate a corresponding noisy image. We then use these noisy images to assess the RDN network and the resultant denoised test images to validate the SSD object detector.

B. Implementation Details

In the RDN stage of our network, we use the kernel size 3 × 3 for every convolutional layer in the network. However, for the local feature and dense feature fusion, we use 1 × 1 as the filter size. We keep the input size fixed throughout the network by padding zeros to input each side. We take G_0 = 64 filters for existing feature fusion layers, and G = 64 for other Conv layers followed by ReLU [32]. As for the network, we use C = 5 Conv layers in the RDB block and D = 10 number of RDBs. When training the model, we employ the settings of [34], according to which, we select 16 random low-quality RGB patches with 48 ×48 as the input size, in each training batch. We implement random image augmentation to the patches, including horizontal or vertical flip and 90° rotation. We used the Adam Optimizer [37] to train the denoising model for 50 epochs with input batch size of 16. We initially set the learning rate to be 1×10^{-4}, and it decayed half for every 10 epochs. We train the RDN network

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:32:36 UTC from IEEE Xplore. Restrictions apply.

Method	σ	mAP	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	mbike	person	plant	sheep	sofa	train	tv
SSD-1	n/a	74.3	76.5	79.0	70.9	65.5	52.1	83.1	84.7	86.4	52.0	81.9	65.7	84.8	84.6	77.5	76.7	38.8	73.6	73.9	83.0	72.6
SSD-2	n/a	79.6	80.9	86.3	79.0	76.2	57.6	87.3	88.2	88.6	60.5	85.4	76.7	87.5	89.2	84.5	81.4	55.0	81.9	81.5	85.9	78.9
SSD-1	11	19.7	24.2	23.4	15.1	11.9	11.3	21.4	28.4	29.8	10.9	12.7	19.6	24.4	27.1	17.5	25.9	12.1	26.6	13.8	18.6	20.1
SSD-1	16	19.4	24.5	24.2	14.7	11.8	11.0	18.9	28.1	29.2	10.5	13.0	18.3	24.1	26.2	18.3	25.6	12.6	24.6	13.2	19.2	19.6
SSD-1	20	19.9	27.5	25.7	15.2	11.2	10.8	22.2	28.7	28.7	10.8	12.0	19.7	23.3	29.3	17.4	25.7	12.1	24.8	14.1	19.0	19.3
SSD-1	23	19.8	25.3	25.3	15.0	11.8	10.6	21.3	29.1	28.8	10.8	12.8	19.2	24.2	25.3	18.6	26.6	12.8	24.6	13.6	20.0	20.9
SSD-1	26	20.2	23.0	26.8	16.9	11.5	11.1	20.4	29.4	30.1	11.0	13.2	20.8	24.6	26.5	20.2	26.3	13.3	25.4	13.8	18.6	20.0
Ours	11	81.6	88.4	82.6	79.2	72.3	56.8	88.3	84.7	90.9	67.7	86.1	78.4	90.3	89.7	87.5	82.2	66.3	79.5	86.1	90.4	83.9
Ours	16	81.3	87.2	82.6	79.0	70.7	56.6	88.3	84.5	90.8	66.6	86.4	78.9	90.5	89.5	86.1	82.1	65.6	79.9	85.7	90.4	83.9
Ours	20	80.9	87.4	82.9	78.3	70.5	56.4	88.2	84.2	90.9	66.7	84.6	78.2	90.4	89.3	85.6	82.0	64.0	79.4	85.4	90.2	83.8
Ours	23	80.7	87.5	82.8	77.7	70.0	55.9	87.7	83.6	90.8	66.6	84.3	78.7	90.3	89.3	86.0	81.4	63.8	79.7	85.5	90.0	83.5
Ours	26	80.1	86.4	82.4	77.3	68.4	54.6	87.1	83.8	90.7	65.9	83.3	78.1	90.1	89.4	85.1	81.0	63.6	77.0	84.9	90.1	82.7

TABLE I: **Detection results of models on PASCAL VOC2007 `test` set**: The first two SSD models are the original method by Lin *et al.* [4]. The SSD-1 model is trained on the combination of VOC2007 and 2012 `trainval` set. And the other one was trained on MS COCO [32], [34] dataset, but additionally, it was fine-tuned on the same data as first. The σ here signifies the variance of noise in the image for evaluation.

for different noise levels, $\sigma = 11, 16, 20, 23, 26$. For the SSD detector part, we resize the denoised image to 300×300 before feeding in the network. The kernel sizes and all other layer parameters are the same as used by Liu *et al.* in [4]. Following their learning and other settings, we initially set the learning rate to 10^{-3}, and weight decay 0.0005. We trained the SSD model with Adam optimizer with 1^{st} moment exponential decay rate 0.9 and batch size 32.

We trained both our models on the Nvidia 2070-SUPER GPU. After training both the networks independently, we use the trained models in a single pipeline for the task of evaluation and testing. For each noise level, we evaluated our entire system and achieved improved results. Code for the paper can be found at: https://github.com/kushagrayadv/ir-with-ssd.

C. Results

Detection Results. We contrast our model with previous SSD methods for the object detection task. These methods did not consider the impact of noise on their performance. We choose the SSD300, the model with 300×300 input size, as the base method for a fair comparison with our model as we used the same method for the detection in our pipeline. We compare with two variants of SSD300. The first one is the model that was trained on the combined PASCAL VOC2007 and 2012 `trainval` set, and we denote this method as SSD-1. And the other SSD300 model was trained on the MS COCO dataset, and was additionally fine-tuned on the same data as the first. We denote this second model as SSD-2. According to the results published in [4], SSD-2 reported 79.6% mAP, significantly than the SSD-1 74.3% mAP. The reason for this increase in the mAP was the availability of additional data for training.

Table 1 represents the results of our method as compared with some earlier detection methods. Our model, for all the noise levels, outperforms both the previous SSD300 models. Our approach shows the best mAP when the noise level is 11. It exhibits an 81.6% mAP, an improvement of 2% when compared to the SSD-2 model, and an increase of 7.3% when compared with the SSD-1 model.

Similarly, our other models also outperform the SSD-1 method with a significant margin in mAP percentage. They also show a slight mAP increase over the SSD-2 model. For $\sigma = 16$, our model displays a rise of 1.7% and 7% over the SSD-2 and SSD-1 methods, respectively. For $\sigma = 20$, a growth of 1.3% over the SSD-2 model and 6.6% over the SSD-1 is achieved. And our model obtains a rise of 6.4% and 1.1% over the SSD-1 and SSD-2, respectively for $\sigma = 23$. Our approach finally increased the mAP by 5.8% when comparing SSD1 and by 0.5% when comparing with SSD-2, in case when $\sigma = 26$. However, there is a reduction in the overall mAP for increasing input noise levels. The reason is the inability of the denoising network to remove the noise from the input completely for higher degree of noise.

We also tested the SSD-1 model on the same noisy image sets we used to evaluate our proposed method. We exhibit the results obtained in table 1. SSD-1 performs poorly on the noisy images and shows approximately 20% mAP for all the noise levels. Thus, it confirms the need for image denoising before object detection in case of noisy images.

	$\sigma = 11$	$\sigma = 16$	$\sigma = 20$	$\sigma = 23$	$\sigma = 26$
BF[36]	29.1	27.2	27.3	27.1	26.7
RME[37]	28.2	27.7	27.2	26.8	26.5
Ours	35.7	33.8	33.9	33.6	33.4

TABLE II: Image denoising results (PSNR) on PASCAL VOC `test` set.

Denoising Results. In table 2, we report the PSNR results of the denoising network in our model on the PASCAL VOC2007 `test` set. We compare it with other methods that have used image denoising as a preprocessing step in object detection [38]. These methods use bilateral filtering (BF) [39] and regularity measure estimation (RME) [38] as the image denoising algorithm. Our approach showed a significant increase in all the noise levels when compared to these earlier methods.

Test images Results. In figure 2, we present the detection results on some test images when processed in our model. We

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:32:36 UTC from IEEE Xplore. Restrictions apply.

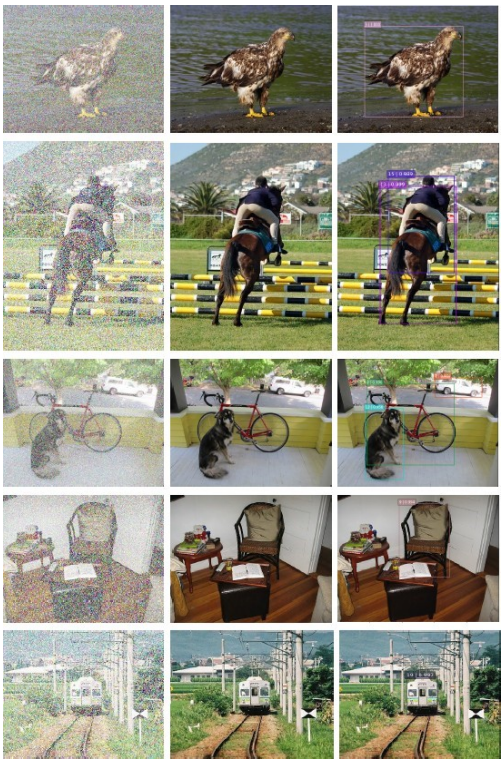


Fig. 2: **Detection results on test images**: The results are on images fused with different AWGN noise levels. The rows represent the images with varying degrees of noise along with their denoised results, and the columns display the inputs and outputs at different stages in our detection pipeline. The 1^{st}, 2^{nd}, 3^{rd}, 4^{th}, and 5^{th} rows show the image results for noise σ = 11, 16, 20, 23, and 26, respectively. The 1^{st} column shows noisy images with noise level designated for each row that we feed in the network. The 2^{nd} column presents the denoised images after the 1^{st} stage of the pipeline, and the 3^{rd} column displays the final output images after object detection.

also conducted some other experiments. We took some test images and estimated noise present in the images. Then we processed it through the trained model with the noise level closest to that of the test image noise to get the denoised output, and finally got the detection. Figure 3 shows the results of this experiment.

V. Conclusion and Future Work

In this paper, we suggest a two-stage object detection algorithm, that comprises the image denoising process as the first stage and the principal object detection task as the later stage to improve the previous detection pipelines. Our model implements the Residual Dense Network (RDN) to restore the high-quality image from a noisy one in the image denoising phase. It extracts local features from the densely connected convolutional layers and simultaneously learns hierarchical functions for image denoising. The next stage performs the primary object detection process on the denoised images via



Fig. 3: Results of test images when processed with the images' estimated noise

the well-known SSD detector. The model can be employed in autonomous vehicles, where object detection algorithms are of prime importance to the passengers and the vehicle's safety. We achieved significant results on the Pascal VOC test set. We reached 81.6% mAP for the case when the variance of the noise was 11. For other noise levels, the technique achieved at least 80% mAP.

As the scope for future work, we can implement the denoising and detection method in an end-to-end manner. The image denoising network and detection network would be incorporated in a single convolutional network and would entertain end-to-end training.

References

[1] S. Milani, R. Bernardini, and R. Rinaldo, "Adaptive denoising filtering for object detection applications," in *19th IEEE International Conference on Image Processing*, 2012, pp. 1013–1016.

[2] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 580–587.

[3] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788.

[4] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *Computer Vision – ECCV 2016*, 2016, pp. 21–37.

[5] A. K. Boyat and B. K. Joshi, "A review paper: Noise models in digital image processing," *ArXiv*, vol. abs/1505.03489, 2015.

[6] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[7] W. Dong, L. Zhang, G. Shi, and X. Li, "Nonlocally centralized sparse representation for image restoration," *IEEE Transactions on Image Processing*, vol. 22, no. 4, pp. 1620–1630, 2013.

[8] A. Buades, B. Coll, and J. Morel, "Nonlocal image and movie denoising," *International Journal of Computer Vision*, vol. 76, pp. 123–139, 2008.

[9] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017.

[10] K. Zhang, W. Zuo, and L. Zhang, "Ffdnet: Toward a fast and flexible solution for cnn-based image denoising," *IEEE Transactions on Image Processing*, vol. 27, pp. 4608–4622, 2018.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:32:36 UTC from IEEE Xplore. Restrictions apply.

[11] P. Liu, H. Zhang, K. Zhang, L. Lin, and W. Zuo, "Multi-level wavelet-cnn for image restoration," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 886–88 609, 2018.

[12] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image restoration," *IEEE transactions on pattern analysis and machine intelligence*, 2020.

[13] M. Everingham, S. Eslami, L. Van Gool, C. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes challenge: A retrospective," *International Journal of Computer Vision*, vol. 111, pp. 98–136, 2015.

[14] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91–110, 2004.

[15] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, 2005, pp. 886–893.

[16] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1627–1645, 2010.

[17] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, vol. 1, 2001, pp. I–I.

[18] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, p. 84–90, May 2017. [Online]. Available: https://doi.org/10.1145/3065386

[19] J. Deng, W. Dong, R. Socher, L. Li, Kai Li, and Li Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[20] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision (IJCV)*, vol. 115, no. 3, pp. 211–252, 2015.

[21] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, p. 1137–1149, Jun. 2017. [Online]. Available: https://doi.org/10.1109/TPAMI.2016.2577031

[22] R. Girshick, "Fast r-cnn," in *IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1440–1448.

[23] E. Simoncelli and E. H. Adelson, "Noise removal via bayesian wavelet coring," *Proceedings of 3rd IEEE International Conference on Image Processing*, vol. 1, pp. 379–382 vol.1, 1996.

[24] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," in *Proceedings of the 11th Annual International Conference of the Center for Nonlinear Studies on Experimental Mathematics*, 1992, p. 259–268.

[25] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.

[26] A. Buades, B. Coll, and J. Morel, "A non-local algorithm for image denoising," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 2, 2005, pp. 60–65.

[27] V. Jain and H. S. Seung, "Natural image denoising with convolutional networks," in *Proceedings of the 21st International Conference on Neural Information Processing Systems*, ser. NIPS'08, 2008, p. 769–776.

[28] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *ArXiv*, vol. abs/1502.03167, 2015.

[29] K. Zhang, W. Zuo, and L. Zhang, "Learning a single convolutional super-resolution network for multiple degradations," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3262–3271, 2018.

[30] Y. Tai, J. Yang, X. Liu, and C. Xu, "Memnet: A persistent memory network for image restoration," *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 4549–4557, 2017.

[31] T. Tong, G. Li, X. Liu, and Q. Gao, "Image super-resolution using dense skip connections," in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 4809–4817.

[32] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, vol. 15, 2011, pp. 315–323.

[33] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2015.

[34] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2017, pp. 1132–1140.

[35] PASCAL VOC 2012, "The pascal visual object classes challenge 2012," http://host.robots.ox.ac.uk/pascal/VOC/voc2012/.

[36] R. Timofte, E. Agustsson, L. V. Gool, M. Yang, L. Zhang, B. Lim, S. Son, H. Kim, S. Nah, and K. M. L. *et al.* , "Ntire 2017 challenge on single image super-resolution: Methods and results," in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2017, pp. 1110–1121.

[37] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *CoRR*, vol. abs/1412.6980, 2015.

[38] S. A. Milyaev and I. Laptev, "Towards reliable object detection in noisy images," *Pattern Recognition and Image Analysis*, vol. 27, pp. 713–722, 2017.

[39] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *6th International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, 1998, pp. 839–846.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:32:36 UTC from IEEE Xplore. Restrictions apply.

Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Impact of interface trap charges on Junctionless double and triple metal gate High-k Gate All Around Nanowire FET based Alzheimer Biosensor

Rishu Chaujar (chaujar.rishu@dtu.ac.in)
Delhi technological University https://orcid.org/0000-0002-0161-8449
Mekonnen Getnet Yirak
Delhi Technological University

Research Article

Keywords: Interface trap charge, bio-sensor, junctionless-TMG-High-k, gate all around, NWFET, atlas-3D

DOI: https://doi.org/10.21203/rs.3.rs-504559/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Impact of interface trap charges on Junctionless double and triple metal gate High-k Gate All Around Nanowire FET based Alzheimer Biosensor

Rishu Chaujar [a,*] Mekonnen Getnet Yirak[ab],

[a] Applied Physics Department, Delhi Technological University, Delhi, INDIA

[b]Physics Department, Debre Tabor University, Debre Tabor, Ethiopia

*Corsponding author email: chaujar.rishu@dtu.ac.in

chaujar.rishu@dtu.ac.in, mekonnengetnet01@gmail.com

Abstract

In this work, junctionless double and triple metal gate high-k gate all around nanowire field-effect transistor-based APTES biosensor has been developed to study the impact of ITCs on device sensitivity. The analytical results were authenticated using ''ATLAS-3D'' device simulation tool. Effect of different interface trap charge on the output characteristics of double and triple metal gate high-k gate all around junctionless NWFET biosensor was studied. Output characteristics, like transconductance, output conductance,drain current, threshold voltage, subthreshold voltage and switching ratio, including APTES biomolecule, have been studied in both devices. 184% improvement has been investigated in shifting threshold voltage in a triple metal gate compared to a double metal gate when APTES biomolecule immobilizes on the nanogap cavity region under negative ITCs. Based on this finding, drain off-current ratio and shifting threshold voltage were considered as sensing metrics when APTES biomolecule immobilizes in the nanogap cavity under negative ITCs which is significant for Alzheimer's disease detection. We signifies a negative ITC has a positive impact on our proposed biosensor device compared to positive and neutral ITCs.

Keywords: - Interface trap charge; bio-sensor; junctionless-TMG-High-k; gate all around; NWFET; atlas-3D

1. INTRODUCTION

Recently, the primary research interest area is biomolecular species detection; detection of these biomolecular species is used for early detection of biological diseases, like Alzheimer's, breast cancer, ovarian cancer (viral diseases), hepatocellular carcinoma (HCC), and other biological diseases. Biological molecules such as RNA, DNA, Uricase, APTES, Streptavidin, ChO_X, and Biotin play a vital role in the screening of those diseases[1][2][3]. However, biomolecule detection using a conventional method like surface plasmon resonance, optical measurements, spectrometry, and quartz crystal microbalances are involved multi-stage processes, expensive, time-consuming, and unsuitable for on-line intensive care[4][5][6][7]. To overcome these drawbacks, the new type of FET like, Field-effect transistors (FET) immunosensors[8], Tunnel filed effect transistor[9], and Junctionless FinFET[10] has emerged; because of easy scalability, low power consumption, label-free and high packing density for the detection of biomolecules[1][2].

As the technology grown, FET has been scaled down to increase transistors in a single chip. Bulk metal-oxide field-effect transistors (MOSFETs) performance enhancement has been reached saturation at 20/22nm

complementary metal-oxide-semiconductor generation, forcing Intel's to shift from bulk to multi-gate device structure to continue on the track of Moor's law[11][12]. However, Leakage current, mobility degradation, subthreshold swing, short response time, hot carrier, power consumption, parasitic resistance, and threshold voltage roll-off have been a serious problem when the device shifts to ultra-small scale device dimension[13][14]. Moreover, the transistor faces other significant challenges at ultra-small scale such as electromigration, thermal budget management, 3-D process integration, and dependence of defects along with the oxide interface trap charge due to high electric field[15][16][17]. Gate leakage current limit's device reliability and memory even lower than simple transistor logic under thinner gate oxide, like SiO_2)[18] [19]. To overcome those device fabrication problems, different approaches like multi-gate device architecture[20] and a dielectric modulating technique (replacing thinner silicon oxide (SiO_2) gate insulator with high-k dielectric constant material)[21] have been proposed. High-k gate oxide makes thicker equivalent oxide thickness (EOT), and it is used to over com quantum mechanical tunneling and SCEs under the nanoscale regime of gate oxide thickness[13][22][23]; this can be obtained using **eq. (1)**[13].

$$\mathrm{EOT} = \mathrm{t}_{\mathrm{SiO}_2} + \left(\frac{\mathrm{K}_{\mathrm{SiO}_2}}{\mathrm{K}_{\mathrm{high-k}}}\right)\mathrm{t}_{\mathrm{high-k}} \qquad (1)$$

Where $\mathrm{t}_{\mathrm{SiO}_2}$ and $\mathrm{t}_{\mathrm{high-k}}$ are the thickness of SiO_2 and HfO_2 dielectric material respectively, $\mathrm{K}_{\mathrm{SiO}_2}$and K_{HfO_2} are dielectric constants of $SiO_2$and HfO_2 oxide material, respectively.

Another problem of CMOS transistor is p-n junction related problem(random dopant fluctuations and an abrupt(unexpected)junction), when channel length and thickness are below 10nm[16][24], which requires expensive fabrication techniques due to diffusion of impurities between p- or n-type drain/source region and n- or p-type body region leads to great difficulty in the fabrication process of small scale devices[25][26]. Dopingless (DL) transistors having uniform doping profiles from source to drain through the channel region are considered significant device architecture for mass production of a microchip in the semiconductor industry. They will have a solution for an abrupt junction problem [24][27]. Dopingless transistor has various advantages over CMOS transistors, including reducing SCEs and accessible for the fabrication process[28]. Therefore, one of the most promising candidate architectures for CMOS fabrication technology is dopingless (DL) MOSFET because it is cost-effective, easy to fabricate, and low power consumption[29]. The impact of fringing capacitance from the polysilicon gate can be controlled using dissimilar gate (multigate)/different gate electrodes with different work-function[30].

Additionally, the ultrathin channel material favors doping less or lightly doped to eradicate threshold-voltage fluctuation due to the variation of dopant atoms and mobility degradation due to depletion charges fluctuation[31]. Nevertheless, in nanoscale MOSFETs, induced hot-carrier, resulting from impact ionization in the channel near the drain junction, becomes a significant reliability concern. In general, hot carriers are injected into the gate oxide due to the high drain electric field, which gives rise to localized interface states and oxide charges near the drain junction and degrades the electrical characteristic parameters [32].

In this paper, effect of interface trap charges (ITCs) on the electrical output characteristics of triple metal gate high-k gate all around junctionless NWFET based Alzheimer biosensor has been proposed using atlas-3D device simulation tool to study the effect of ITCs on device performance; considering trap charges near the insulator/semiconductor interface as shown in **Fig 1**. Including APTES biomolecule on output characteristics of n-type double and triple metal gate JL-GS-GAA-NWFET, has been analyzed under room

temperature. The outcome of localized/interface trap charge on the device output physical characteristics, such as output resistance, threshold voltage variation, transconductance, early voltage, intrinsic voltage gain, output conductance, device efficiency, and surface potential, has been extensively studied.

1.2 DEVICE SIMULATION AND STRUCTURE

Device structure for n-type junctionless high-k triple metal gate all around NWFET based biosensor with ITCs have been illustrated in **Fig 1(b)**. Here, $L_1(6nm)$, $L_2(8nm)$, $L_3(6nm)$ are the lengths of M_1, M_2, and M_3 respectively and $L_4(10nm)$, are length of the nanogap cavity and silicon dioxide (SiO_2), etching near to drain and source end, and L (20 nm) channel length. T_1, T_2, $T_3$and T_4 are the thickness of metal gate, hafnium oxide, nanogap cavity, and interface (SiO_2) oxide, respectively, and 2R is the diameter of the channel. A (T_4 =0.3nm) thickness of SiO_2 interface layer is considered between hafnium oxide and silicon film to create a compatible region between hafnium oxide and silicon film for carrier mobility[33]. The three gate materials denoted by $M_1, M_2$and M_3have different work-function given by $\Phi_{M1} = 4.86$, $\Phi_{M2} = 4.96$ and $\Phi_{M3} = 4.50$, respectively. The intermediate work-function (Φ_{M1}) near to the source is used to control electron saturation velocity and doping fluctuation, the highest work function(Φ_{M2}) between source end and drain end is used to control potential profile along the channel, which intern reduce drain induced barrier lowering (DIBL) and lower work-function near to drain(Φ_{M3}) is used for screening effects and controls high drain electric field fluctuation due to V_{DS}[34][35]. The nanogap cavity region used as a detecting site in which the target biomolecules (APTES, ϵ =3.57) and (air= ϵ =1.0) are assumed to be uniformly immobilized in this region by introducing their dielectric constant[36]. Different types of ITCs (positive and negative) given by ($N_f = \pm 5x10^{16} cm^{-2}$)and also without interface (Neutral) charge ($N_f = 0$) are simulated interchangeably to examine the ITCs impact of the proposed device performance.

Parameters	DG-GAA-JL-NWFET		TG-GAA-JL-NWFET		
Length of the channel (nm)	20		20		
Hafnium Oxide thickness (nm)	HfO_2=1.5 &SiO_2=0.3		HfO_2=1.5 &SiO_2=0.3		
Oxide thickness near to the source (nm)	SiO_2=1		SiO_2=1		
Oxide length near to the source (nm)	SiO_2=10		SiO_2=10		
Nanogap cavity length near to drain (nm)	10		10		
Drain/Source thickness (nm)	10		10		
Source/Drain length (nm)	10		10		
Nanogap cavity thickness (nm)	1		1		
The diameter of silicon (nm)	10		10		
Interface trap charges (ITCs)	$\pm 5x10^{12} cm^{-2}$		$\pm 5x10^{12} cm^{-2}$		
Source/Drain & Channel Doping (N_D+)	$10^{19} cm^{-3}$		$10^{19} cm^{-3}$		
Oxide Dielectric constant	SiO_2=3.9&HfO_2=25.0		SiO_2=3.9&HfO_2=25.0		
Gate Work functions (eV)	4.50	4.86	4.86	4.96	4.50

Table I:- Proposed device structural parameters.

Several simulations have been carried out using the atlas-3D device simulator to describe the proposed device's electrical properties based on these device descriptions. Here different models are applied, such as (CONMOB) concentration-dependent mobility [37] model, to study doping versus mobility distribution along with bandgap narrowing (BGN) model. Shockley–Read–Hall (SRH) model and Boltz-man transport equation account for minority carriers[38][39]. (CCSMOB) Carrier-carrier scattering mobility model used to study carrier mobility and concentration. CVT models are also considered to study perpendicular and parallel field-dependent carrier mobility and mobility roll-off [38][40]. Carrier transport equations have been solved using Newton's and Gummel's numerical methods. Quantum mechanical models have not been invoked in this simulation because the silicon film radius is greater than 4nm[10]. The impression of interface trap charge on device performance has been studied by considering uniformly distributed interface trap charge ($N_f = \pm 5x10^{12} cm^{-2}$) at the Si–$SiO_2$ interface, as illustrated in **Fig 1(b).** The reason is that interface trap charge near the drain side is vulnerable to high drain electric field for under short channel device[31]; this high electric field induces hot carriers and leads to rising localized charge at the Si–SiO_2 interface and will create permanent damage to the device performance[41].

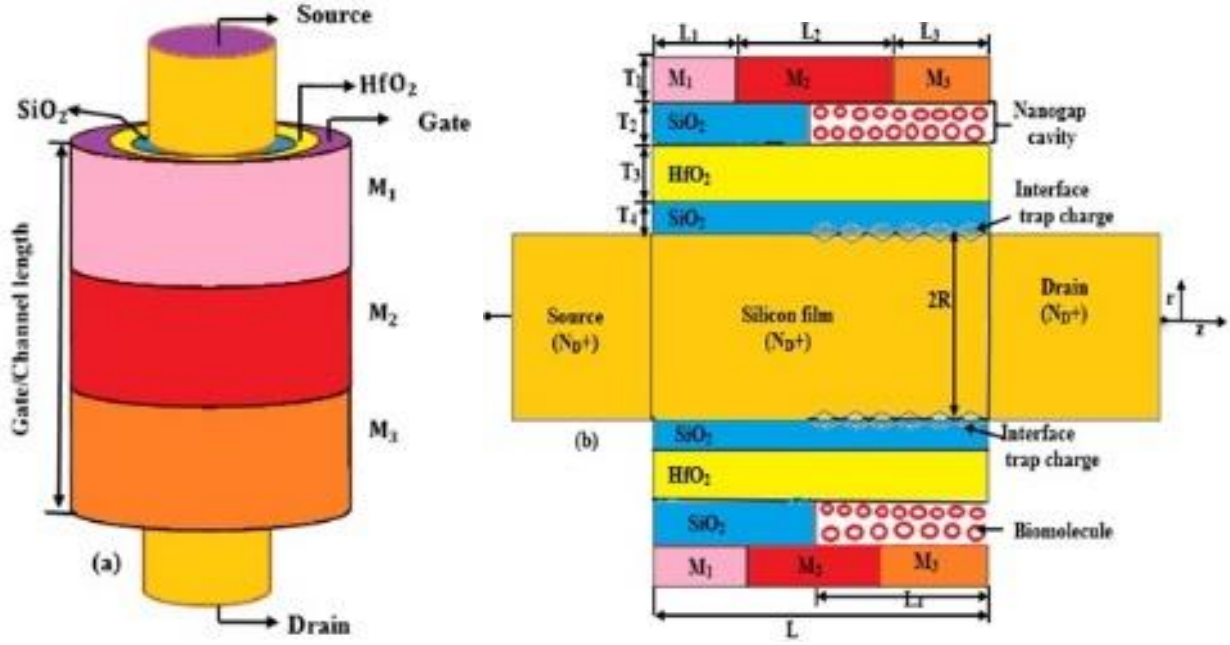


Fig 1. (a) 3D schematic diagrame and **(b)** 2D cross-sectional interpretation with nanogap cavity for n-type triple metal gate high-k gate AA-JL-NWFET.

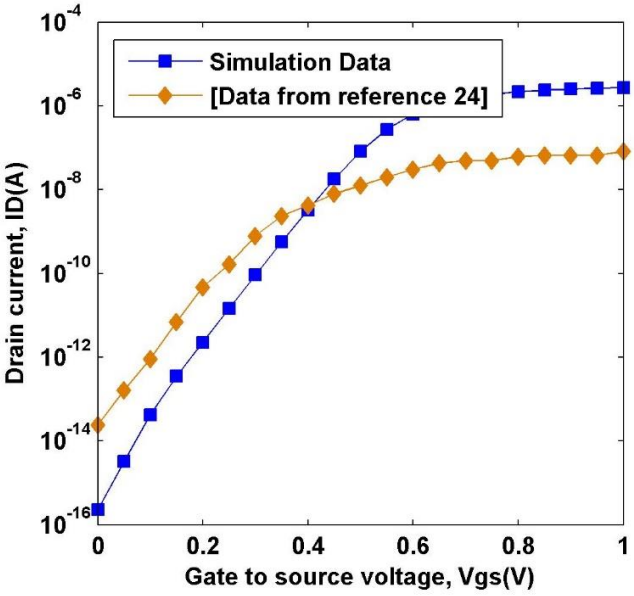


Fig 2. Calibration of simulation results compared with data from reference [24]for n-type triple metal gate high-k gate AA-JL-NWFET without ITCs and nanogap cavity region at ($V_{DS} = 0.05V$).

1.3 RESULTS AND DISCUSSION

The output characteristics of double and triple metal gate junctionless high-k gate all around NWFET-based Alzheimer Biosensor have been examined incorporating interface trap charges.

a) Effect of ITCs on drain current

Fig 3. Illustrates the simulated results of I_D-V_G transfer characteristics on log scale at $V_D = 0.05V$ with interface trap charge of double and triple metal gate. The triple metal gate has a lower leakage current compared to a double metal gate. The reason is that the intermediate work-function (Φ_1) near to the source region is used to control electron saturation velocity and hot carrier effects through the channel. Highest work-function (Φ_2) lowers drain-induced barrier lowering and parasitic resistance, thereby increases gate controllability[15]. Since ITCs rise at $S - SiO_2$ interface, it will accept an electron for an acceptor-type ITC, and it acts as a fixed negative interface trap charge. Likewise, the donor type interface trap acts as a positive localized charge. These interface traps can be transferred into equivalent interface fixed charge led to band bending under the gate. Greater change in flat band voltage under the damaged region affects electron mobility in the channel[41]. **Fig 3.** Illustrates lower leakage current for negative interface charge and higher leakage current for positive interface charge in both devices; this is because negative ITC mainly reduces DIBL and hot carrier effects in contrary increases mobility carries in the channel. For triple metal gate,I_{OFF} results of positive and negative interface trap charges are $1.5x10^{-15}$A and $7.97x10^{-17}$A respectively at $V_{GS} = 0$ and $V_{DS} = 0.05V$.

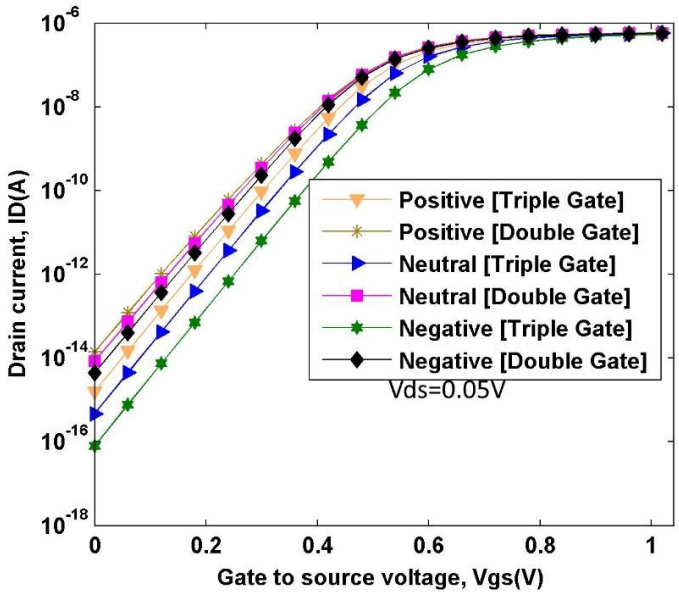


Fig 3. Effect of interface trap charge on I_D-V_G for n-type double and triple metal gate high-k gate AA-JL-NWFET without cavity region.

b) Effect of ITCs on subthreshold swing and threshold voltage (V_{th})

Fig 4. Illustrates simulated results of **(a)** threshold voltage (V_{th}) **(b)** subthreshold swing at $V_{gs} = 1.0V$ and $V_{DS} = 0.05V$ with ITCs for double and triple metal gate. A lower subthreshold slope has been observed under the triple metal gate compared to the double metal gate. The work-function (Φ_1) near to the source region is used to control carrier injection through the channel and the highest work-function the (Φ_2) develops a barrier potential used to reduces reverse current from the drain through the channel, thereby increasing mobility carriers [15]. Lowest work-function of the metal gate (Φ_3) drain side reduces hot-carriers induced due to high electric field. Variation in threshold voltage is due to different localized/interface trap charges, which introduces a step potential profile in the channel due to the high electric field near the drain side; will rise gate and substrate current; this creates damage in the oxide interface near the drain junction. The metal gate work-function (Φ_3) lowers the peak electric field at the drain side and raises the average electric field near the gate to enhance device performance[41]. It is clear that **Fig 4**. Demonstrates a negative interface trap charge increase threshold voltage while decreasing under positive ITC due to induced hot carriers in the case of positive ITC; negative ITC increases carrier concentration on the channel for n-type channel CMOS device reduces leakage current and subthreshold swing[42]. Also, the fixed localized charge at the interface cause band bending under the gate, leading to changes in flat band voltage in the damaged region led to a change threshold voltage[43]. Higher degradation of SS in the triple metal gate for negative ITCs is investigated due to higher electron mobility and threshold voltage compared to positive interface trap charge; this enhances gate electrostatic control[44], as shown in ***Fig* 4.(b)** compared to a double metal gate transistor.

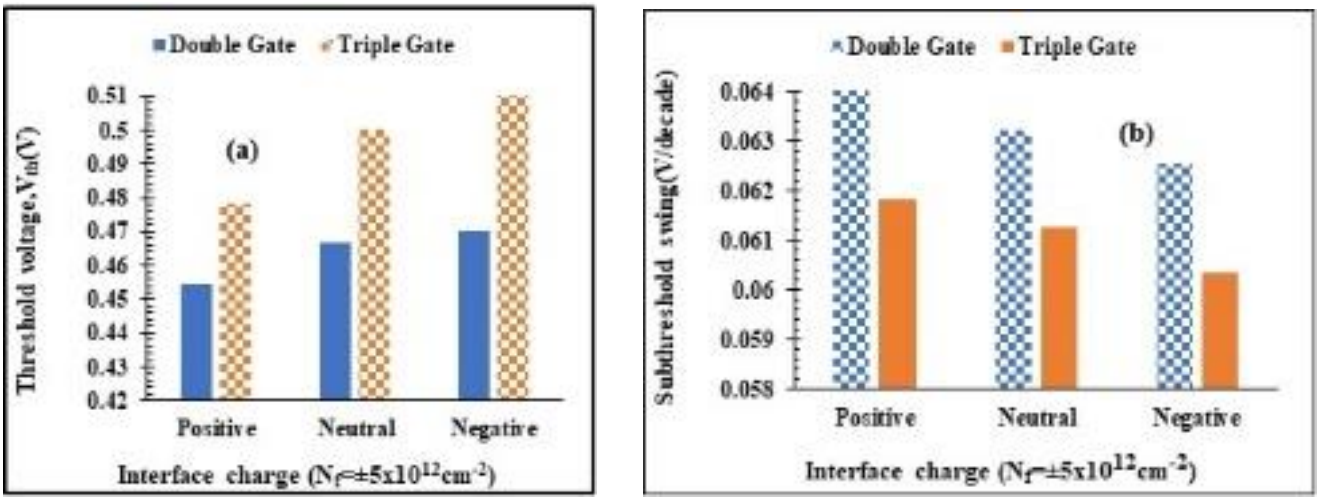


Fig 4. Variation of **(a)** V_{th} and **(b)** subthreshold swing, with ITCs for n-type double and triple metal gate high-k gate AA-JL- NWFET without nanogap cavity at ($V_{DS} = 0.05V \; and \; V_{GS} = 1.0V$).

c) Effect of interface charge on device sensitivity (V_{th} and $S_{I_{OFF}}$)

The threshold voltage is an essential characteristic of NW-MOSFET biosensor used as the sensing parameter to detect the device's sensitivity when the biomolecule interacts with the device. The detection mechanism of biomolecule species is carried out by introducing APTES (biomolecule) dielectric constant in the nanogap cavity. **Fig 5**. **(a)** demonstrates the effect of APTES biomolecule on shifting threshold voltage (ΔV_{th}) and **(b)** drain off-current ratio for n-type double and triple metal gate JL-high-k-GAA-NWFET with ITCs. Shifting threshold voltage (V_{th}) of MOSFET device is obtained using **eq. (2)**[45].

$$\Delta V_{th} = |V_{th}(\epsilon_{APTES} = 3.57) - V_{th}(\epsilon_{air} = 1)| \qquad (2)$$

Where ($\epsilon_{APTES} = 3.57$) represents APTES biomolecule dielectric constant and ($\epsilon_{air} = 1$) is air dielectric constant placed in the nanogap cavity region.

Figure 5(a). Demonstrates higher threshold voltage shifting for negative interface charge in both devices because negative ITC increases mobility carrier throughout the channel and reduce threshold voltage roll-off. This is because negative ITC is mainly used to improve mobility carriers and stabilize high electric fields near the drain for the n-type channel. For a positive ITCs, the device exposed to short channel effects led to threshold voltage decay. In addition to the interface trap charge, biomolecule (APTES) affects device sensitivity (shifting threshold voltage), as illustrated in **Fig 5(a)**.

Fig 5(b). Shows drain off-current ratio ($S_{I_{off}}$) is decrease for negative interface charge due to lower variation of leakage current. The result shows that higher drain off-current ratio for positive and neutral compared to negative ITC; because of, hot-carries induced by localized charge and high drain electric field near drain side in the presence of positive and neutral ITCs increase leakage current than negative ITCs; also, drain induced barrier lowering effects experienced by positive and neutral ITCs. At negative ITCs, with and without biomolecule, I_{OFF} ratio is lower than positive ITCs due to less induced hot-carriers and average electric field experienced by negative interface charge. Sensitivity ($S_{I_{off}}$) or drain-off current ratio is given by **eq. (3)**.

$$S_{I_{off}} = \frac{I_{off}(\text{with biomolecule Species})}{I_{off}(\text{without biomolecule Species})} \Big| \text{at } V_{gs} = 0 \qquad (3)$$

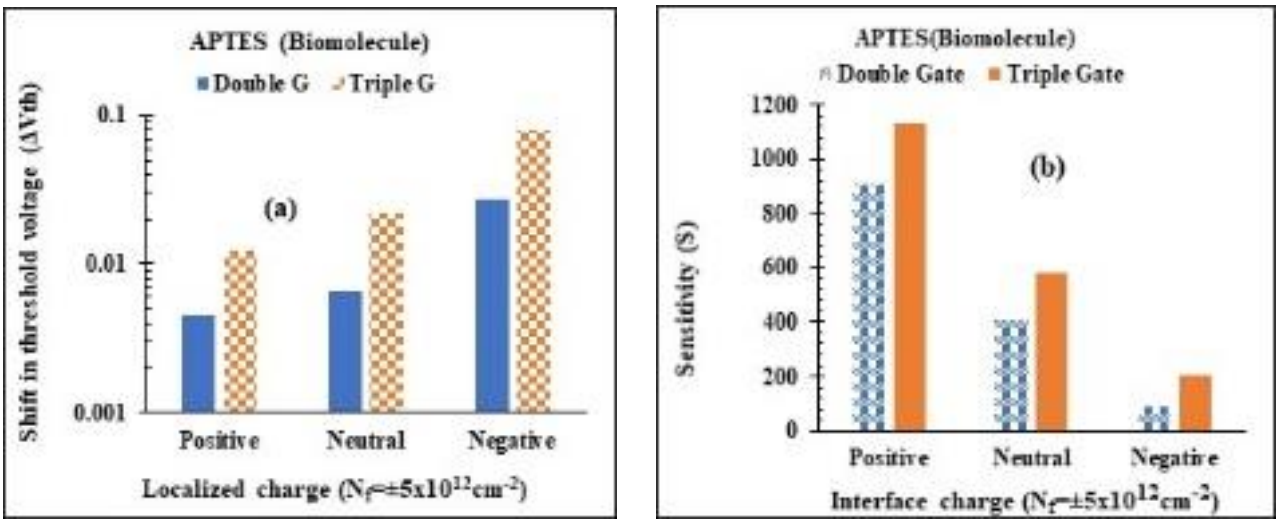


Fig 5. Variation of **(a)** shifting threshold voltage **(b)** sensitivity ($S_{I_{off}}$), with ITCs for n-type double and triple metal gate high-k gate all around JL-NWFET at ($V_{DS} = 0.05V$ and $V_{GS} = 1.0V$).

Generally, Higher variation of threshold voltage is observed in a triple metal gate than a double metal gate. The work-function (Φ_1) near the source region increases uniformity, and mobility of carriers through the channel reduces hot carrier and impact ionization effects. Negative ITC in both devices is used to control electron velocity through the channel. In triple metal gate, the intermediate higher work-function the (Φ_2) and lower work-function(Φ_3) reduces threshold voltage roll-off and drain-induced barrier lowering impacts in addition to parasitic resistance, thereby enhance gate capacity[42].

d) Impact of ITCs on switching ratio and leakage current

Fig 6. Illustrates the variation of **(a)** switching ratio $\left(\frac{I_{on}}{I_{off}}\right)$, **(b)** leakage current at $V_D = 0.05V$ with ITCs ($N_f = \pm 5 \times 10^{12}$) in both double and triple metal gate transistors. Lower leakage current and higher switching ratios are investigated in triple metal gate than double metal gate device. This is the dissimilar gate electrodes effect that reduces gate parasitic resistance and induced hot-carriers caused by the polysilicon gate and high drain electric field. In both devices, higher switching and lower leakage current are investigated with negative ITC compared to positive and neutral ITCs; negative ITCs can mitigate DIBL and induce hot-carrier effects, thereby increasing mobility carriers [42]. Thus, higher switching (I_{ON}/I_{OFF}) ratio in the case of negative interface trap charges while reduced switching (I_{ON}/I_{OFF}) ratio in the case of neutral and positive interface charges as illustrated in **Fig 6**(a).

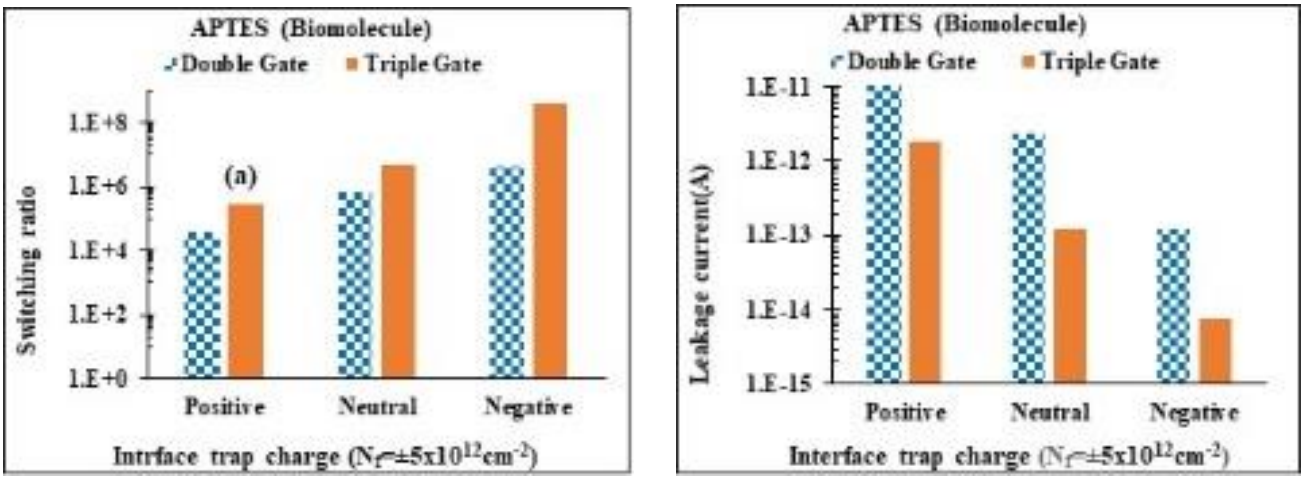


Fig 6 (a). Variation of switching ratio **(b)** leakage current, with different ITCs for n-type double and triple metal gate high-k gate all around junctionless NWFET at ($V_{DS} = 0.05V \; and \; V_{GS} = 1.0V$).

When APTES (biomolecule) immobilizes in the nanogap cavity, drain on current increases while drain off-current decreases. The reason is higher biomolecule dielectric constant decreases electron tunneling from the gate and body substrate, led to improvs gate electrostatic control in ultra-small gate oxide thickness in short channel device [14]. Drain off current ratio decreases/increases for positive /negative interface trap charge in all operation regions. Reduced I_{on}/I_{off} ratio for positive interface trap charge and improved I_{on}/I_{off} ratio for negative interface trap charge, as shown in **Fig 6(a)**. For instance, switching (I_{ON}/I_{OFF}) ratio of APTES (biomolecule) are $4.18x10^{6}$ and $4.18x10^{8}$ respectively for double and triple metal gates of the proposed device at negative interface trap charge (ITC).

Fig 6(b). Demonstrates the impact of ITCs on leakage current with APTES biomolecule for double and triple metal gate high-k gate all around junctionless NWFET. It clearly shows, lower leakage current for negative ITCs than positive ITCs; the reason is that the negative interface trap charge increases mobility carriers through the channel while the positive ITCs reduce mobility carriers through the channel due to trapping localized charge and induced hot-carriers. Reduced charge carriers in the channel due to positive ITC allow the depletion layer to become thicker across the reverse junction that allows the flow of charge carriers in OFF-state, resulting from higher leakage current[16]. Hot-carriers generated high drain electric fields, grows serious problems, such as reliability and device efficiency[46].

The output characteristics such as transconductance, and output conductance are known as the device analog parameters, which can be used to examine the device's gain (amplification) and reliability[44].Fig **7.** Illustrates the impact of ITCs on transconductance and output conductance for double and triple metal gate transistors. **Fig 7(a).T**ransconductance g_m is higher for negative ITCs in both devices. Large transconductance of the device delivers greater amplification when other factors are kept constant and it can be given by **eq. (4).**

$$g_m = \frac{\partial I_{DS}}{\partial V_{GS}} \qquad (4)$$

Fig 7(b), illustrates the output conductance (g_d) of double and triple metal gate high-k gate all around junctionless NWFET, which is used to characterize its output resistance ($R_{out} = \frac{1}{g_d}$) of the device. In both device, outconductance is decreased for negative but increases for positive and it can be obtained using **eq. (5)** at constant gate voltage (V_{GS}).

$$g_d = \frac{\partial I_{DS}}{\partial V_{DS}} \qquad (5)$$

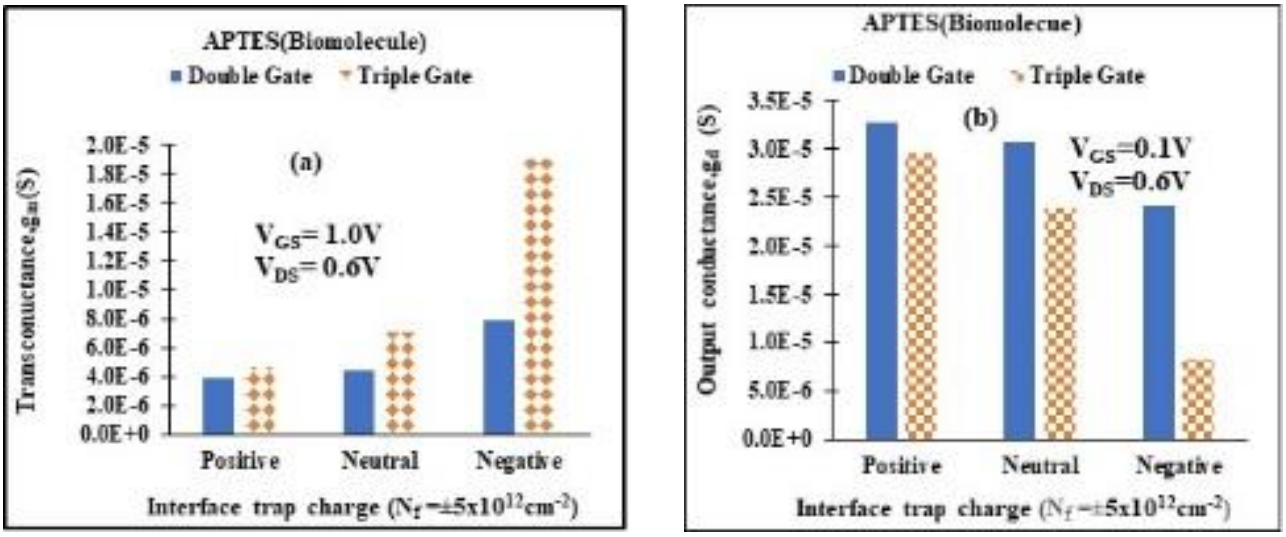


Fig 7. Variation of **(a)** transconductance **(b)** output conductance with ITCs for n-type triple metal gate high-k gate all around junctionless NWFET.

2 CONCLUSION

In this paper, impact of different ITCs on the electrical characteristics of double and TG-high-k-GAA-JL-NWFET based Alzheimer biosensor has been studied. We have observed that ITCs causes the change in switching ratio, shifting threshold voltage, transconductance, output conductance, leakage current, subthreshold slope, in the presence of APTES biomolecule. For instance, the results of SS at negative ITCs are $(62.6 \text{ and } 60.3)\text{mV/decade}$ for double and triple metal gate devices, respectively. Improved output characteristics, such as threshold voltage (0.47V and 0.54V), shifting threshold voltage (2.75mV and 7.81mV), switching ratio (4.18×10^6 and 4.12×10^8) and transconductance (7.88×10^{-6} *and* 1.90×10^{-5}) has been examined at negative ITC of double and triple TG-high-k-GAA-JL-NWFET respectively when APTEs biomolecule immobilizes in the nanogap cavity. Leakage current result when APTES biomolecule interact with the TG-high-k-GAA-JL-NWFET device at nanocavity in the presence of ITCs

($N_f = \pm 5x10^{12} cm^{-2}$) are $1.79x10^{-12}$A, $1.24x10^{-13}$A and $1.79x10^{-15}$A for positive, neutral, and negative ITCs, respectively. Reduced and rising leakage current due to negative and positive ITCs are 93.6% and 1,343.5%, respectively, have been investigated in the proposed device with respect to undamaged region. Increasing and decreasing of surface potential of the proposed device at ITCs ($N_f = \pm 5x10^{12} cm^{-2}$) are 1.54% and 2.39% for positive and negative respectively with respect to undamaged region when APTES biomolecule is immobilized in the nanogap cavity. Generally, we have examined that negative ITC enhances the output characteristics, such as switching ratio, transconductance, output resistance for triple metal than double metal gate -high-k-GAA-JL-NWFET, when APTES biomolecule immobilize in the nanogap cavity. We can conclude that the APTES biosensor with negative ITC is the novel biomarker compared to the positive ITC of the silicon-based device. Since, it introduces reactive amine groups on the silicon surface to detect sugar level variation of saliva caused by Alzheimer's disease. It could also be a potential indicator/biomarker for Alzheimer's disease (AD) detection.

AVAILABILITY OF DATA AND MATERIALS

We announce that the materials mentioned in the manuscript, as well as all related raw data, will be freely accessible to any scientist who wishes to use them for non-commercial purposes without compromising participant confidentiality.

AUTHORS' CONTRIBUTIONS

The main idea of this paper was proposed by Dr. Rishu Chaujar prepared the manuscript initially and performed all the steps of the solutions in this research. Both authors read and approved the final manuscript.

FUNDING

The authors declare that no funding has been received to carry out this research work.

ACKNOWLEDGMENT

The authors are grateful to the Debre Tabor University, Ethiopia, and Delhi Technological University, India, for providing the necessary monetary and material assistance to carry out this research work.

COMPETING INTEREST

The authors declare that they have no competing interest.

DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST

'Not applicable'

RESEARCH INVOLVING HUMAN PARTICIPANTS AND/OR ANIMALS

'Not applicable'

INFORMED CONSENT

'Not applicable'

References

[1] Cibele Gouvea, "Biosensors for Health Applications," in *Open access peer-reviewed chapter*, Pier Andrea Serra, Ed. Universidade Federal de Alfenas, Brazil: IntechOpen.

[2] Y. Wang and G. Li, "Simulation of a silicon nanowire FET biosensor for detecting biotin/streptavidin binding," *2010 10th IEEE Conf. Nanotechnology, NANO 2010*, pp. 1036–1039, 2010.

[3] Y. Chen *et al.*, "Field-Effect Transistor Biosensor for Rapid Detection of Ebola Antigen," *Sci. Rep.*, vol. 7, no. 1, pp. 4–11, 2017.

[4] K. Y. Park, M. S. Kim, and S. Y. Choi, "Fabrication and characteristics of MOSFET protein chip for detection of ribosomal protein," *Biosens. Bioelectron.*, vol. 20, no. 10 SPEC. ISS., pp. 2111–2115, 2015.

[5] A. C. M. de Moraes and L. T. Kubota, "Recent trends in field-effect transistors-based immunosensors," *Chemosensors*, vol. 4, no. 4, 2016.

[6] H. C. Lau, T. E. Bae, H. J. Jang, J. Y. Kwon, W. J. Cho, and J. O. Lim, "Saliva-based screening approach for Alzheimer's disease via the cell-oriented ion-sensitive field-effect transistor," *Sens. Lett.*, vol. 12, no. 6–7, pp. 1096–1101, 2014.

[7] B. Shui *et al.*, "Biosensors for Alzheimer's disease biomarker detection: A review," *Biochimie*, vol. 147, no. January, pp. 13–24, 2018.

[8] Ana Carolina Mazarin de Moraes and Lauro Tatsuo K, "Recent Trends in Field-Effect Transistors-Based Immunosensors," *hemosensors Rev.*, 2016.

[9] J. Madan and R. Chaujar, "Superlattices and Microstructures Palladium Gate All Around - Hetero Dielectric -Tunnel FET based highly sensitive Hydrogen Gas Sensor," *Superlattices Microstruct.*, pp. 1–8, 2016.

[10] A. Chhabra, A. Kumar, and R. Chaujar, "Sub-20 nm GaAs junctionless FinFET for biosensing application," *Vacuum*, vol. 160, no. December 2018, pp. 467–471, 2018.

[11] S. M. Amoroso *et al.*, "Inverse Scaling Trends for Charge-Trapping-Induced Degradation of FinFETs Performance," *IEEE Trans. Electron Devices*, vol. 61, no. 12, pp. 4014–4018, 2014.

[12] J. Wu, Y.-L. Shen, K. Reinhardt, H. Szu, and B. Dong, "A Nanotechnology Enhancement to Moore's Law," *Appl. Comput. Intell. Soft Comput.*, vol. 2013, pp. 1–13, 2013.

[13] K. P. Pradhan, S. K. Mohapatra, P. K. Sahu, and D. K. Behera, "Impact of high-k gate dielectric on analog and RF performance of nanoscale DG-MOSFET," *Microelectronics J.*, vol. 45, no. 2, pp. 144–151, 2014.

[14] N. Thapa and L. Maurya, "Performance advancement of High-K dielectric MOSFET," *Int. J. Innov. Adv. Comput. Sci.*, vol. 3, no. 3, pp. 98–103, 2014.

[15] Sze, "Performance Investigation of Dual Material Gate Stack Schottky-Barrier Source/Drain GAA MOSFET for Analog Applications," *Phys. Semicond. Devices, Environ. Sci. Eng.*, vol. 10, no. December 2013, pp. 739–751, 2014.

[16] N. Kumar and A. Raman, "Performance Assessment of the Charge-Plasma-Based Cylindrical GAA Vertical Nanowire TFET with Impact of Interface Trap Charges," *IEEE Trans. Electron Devices*, vol. 66, no. 10, pp. 4453–4460, 2019.

[17] P. Saha, P. Banerjee, D. K. Dash, and S. K. Sarkar, "Impact of trapped interface charges on short channel characteristics of WFE high-K SOI MOSFET," *Proc. 3rd Int. Conf. 2019 Devices Integr. Circuit, DevIC 2019*, pp. 118–123, 2019.

[18] A. S. Aji and Y. Darma, "Simulation of Charge-Trapping Effect on Floating Gate Si / Ge / Si Quantum Dots MOSFET Memory with High- κ Tunnel Oxide," in *2013 3rd International Conference on Instrumentation, Communications, Information Technology and Biomedical Engineering (ICICI-BME)*, 2013, no. 1, pp. 269–272.

[19] Prerna, "Future MOSFET Devices using High-k (TiO2) Dielectric," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 1, no. 2, pp. 23–28, 2013.

[20] S. L. Tripathi, R. Mishra, and R. A. Mishra, "Multi-Gate Mosfet Structures With High-K Dielectric Materials," *ournal of Electron Devices*, vol. 16, pp. 1388–1394, 2012.

[21] R. Chau *et al.*, "Application of high-κ gate dielectrics and metal gate electrodes to enable silicon and non-silicon logic nanotechnology," *Microelectron. Eng.*, vol. 80, no. SUPPL., pp. 1–6, 2005.

[22] S. K. Swain *et al.*, "Effect of High-K Spacer on the Performance of Non-Uniformly doped DG-MOSFET," *Proc. 3rd Int. Conf. 2019 Devices Integr. Circuit, DevIC 2019*, pp. 510–514, 2019.

[23] V. Kumar, R. Gupta, R. Preet, P. Singh, and R. Vaid, "Performance Analysis of Double Gate n-FinFET Using High-k Dielectric Materials," *Int. J. Innov. Res. Sci. Eng. Technol.*, vol. 5, no. 7, pp. 13242–13249, 2016.

[24] J. P. Colinge *et al.*, "Nanowire transistors without junctions," *Nat. Nanotechnol.*, vol. 5, no. 3, pp. 225–229, 2010.

[25] V. P. Georgiev *et al.*, “Experimental and Simulation Study of Silicon Nanowire Transistors Using Heavily Doped Channels,” *IEEE Trans. Nanotechnol.*, vol. 16, no. 5, pp. 727–735, 2017.

[26] L. Ansari, B. Feldman, G. Fagas, J. P. Colinge, and J. C. Greer, “Simulation of junctionless Si nanowire transistors with 3 nm gate length,” *Appl. Phys. Lett.*, vol. 97, no. 6, 2010.

[27] Z. Huang, A. Gao, S. Chen, Y. Wang, and T. Li, “Highly Sensitive Juntionless Nanowire Transistor Biosensor in Detecting Breast Tumor Marker,” *Proc. IEEE Sensors*, vol. 2018–Octob, pp. 1–4, 2018.

[28] S. Ghosh, A. Chattopadhyay, and S. Tewari, “Optimization of Hetero-Gate-Dielectric Tunnel FET for Label-Free Detection and Identification of Biomolecules,” *IEEE Trans. Electron Devices*, vol. 67, no. 5, pp. 1–8, 2019.

[29] G. Jawade, Y. V. Chavan, and S. G. Wagaj, “Simulation of dual spacer double gate junctionless transistor,” *Proc. - 2nd Int. Conf. Comput. Commun. Control Autom. ICCUBEA 2016*, pp. 1–4, 2017.

[30] F. Ana and Najeeb-ud-din, “Gate Workfunction Engineering for Deep Sub-Micron MOSFET’s: Motivation, Features and Challenges,” *Int. J. Electron. &Communication Technol.*, vol. 2, pp. 29–35, 2011.

[31] E. G. Ioannidis, A. Tsormpatzoglou, D. H. Tassis, C. A. Dimitriadis, G. Ghibaudo, and J. Jomaah, “Effect of localized interface charge on the threshold voltage of short-channel undoped symmetrical double-gate MOSFETs,” *IEEE Trans. Electron Devices*, vol. 58, no. 2, pp. 433–440, 2011.

[32] N. Garg, Y. Pratap, M. Gupta, and S. Kabra, “Analysis of Interface Trap Charges of Double Gate Junctionless Nanowire Transistor (DG-JNT)for Digital Circuit Applications,” *Proc. Int. Conf. 2018 IEEE Electron Device Kolkata Conf. EDKCON 2018*, pp. 563–567, 2018.

[33] M. I. B. Chowdhury, M. J. Islam, M. J. Islam, M. M. Hasan, and S. U. Farwah, “Silvaco TCAD based Analysis of Cylindrical Gate -All-Around FET Having Indium Arsenide as channel and Aluminium Oxide as Gate Dielectrics,” *J. Nanotechnol. its Appl. Eng.*, vol. 1, no. 1, pp. 1–12, 2016.

[34] A. Goel, S. Rewari, S. Verma, and R. S. Gupta, “Dielectric Modulated Triple Metal Gate All Around MOSFET (TMGAA)for DNA Bio-Molecule Detection,” *2018 IEEE Electron Devices Kolkata Conf.*, vol. 1, pp. 337–340, 2019.

[35] N. Bagga and S. Dasgupta, “Surface Potential and Drain Current Analytical Model of Gate All Around Triple Metal TFET,” vol. XX, no. Xx, pp. 1–8, 2017.

[36] Y. Pratap, M. Kumar, S. Kabra, S. Haldar, R. S. Gupta, and M. Gupta, “Analytical modeling of gate-all-around junctionless transistor based biosensors for detection of neutral biomolecule species,” *J. Comput. Electron.*, vol. 17, no. 1, pp. 288–296, 2018.

[37] D. S. Software, *Atlas User ’ s Manual*, no. 408. 2016.

[38] J. S. Galsin, “Transport Phenomena,” in *Solid State Physics*, Elsevier, 2019, pp. 223–242.

[39] J. Madan and R. Chaujar, “Numerical Simulation of N+ Source Pocket PIN-GAA-Tunnel FET: Impact of Interface Trap Charges and Temperature,” *IEEE Trans. Electron Devices*, vol. 64, no. 4, pp. 1482–1488, 2017.

[40] D. Passeri, A. Morozzi, K. Kanxheri, and A. Scorzoni, “Numerical simulation of ISFET structures for biosensing devices with TCAD tools,” *Biomed. Eng. Online*, vol. 14, no. 2, p. S3, 2015.

[41] Y. S. Jean and C. Y. Wu, “The threshold-voltage model of MOSFET devices with localized interface charge,” *IEEE Trans. Electron Devices*, vol. 44, no. 3, pp. 441–447, 1997.

[42] S. M. W. Localized, “A Compact Model for Threshold Voltage of Interface Trapped Charges,” vol. 58, no. 2, pp. 567–571, 2011.

[43] S. K. Mohapatra, K. P. Pradhan, P. K. Sahu, G. S. Pati, and M. R. Kumar, “The effect of interface trapped charges in DMG-S-SOI MOSFET: A perspective study,” *Adv. Nat. Sci. Nanosci. Nanotechnol.*, vol. 5, no. 4, 2014.

[44] J. Yuan and J. C. S. Woo, “A Novel Split-Gate MOSFET Design Realized by a Fully Silicided Gate Process for the Improvement of Transconductance and Output Resistance,” *IEEE ELECTRON DEVICE Lett.*, vol. 26, no. 11, pp. 829–831, 2005.

[45] E. Rahman, A. Shadman, and Q. D. M. Khosru, “Effect of biomolecule position and fi ll in factor on sensitivity of a Dielectric Modulated Double Gate Junctionless MOSFET biosensor,” *Sens. Bio-Sensing Res.*, vol. 13, pp. 49–54, 2017.

[46] A. Kumar, M. M. Tripathi, and R. Chaujar, “Reliability Issues of In2O5Sn Gate Electrode Recessed Channel MOSFET: Impact of Interface Trap Charges and Temperature,” *IEEE Trans. Electron Devices*, vol. 65, no. 3, pp. 860–866, 2018.

2021 2nd International Conference for Emerging Technology (INCET)
Belgaum, India. May 21-23, 2021

Implementation of a Shadow Filter using CFCC

Prakhar Prakash
Department of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
prakharsinha343@gmail.com

Yashodhan Srivastava
Department of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
yashodhansrivastava9@gmail.com

Shubham Kumar
Department of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
kumarshubham525@gmail.com

Alok Kumar Singh
Department of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
aksingh@dce.ac.in

2021 2nd International Conference for Emerging Technology (INCET) | 978-1-7281-7029-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/INCET51464.2021.9456361

***Abstract*: In this paper a new shadow filter implementation has been designed using Current Follower Current Conveyor (CFCC) block and grounded elements. Shadow filter allows for electronically tunable filter parameters by simply varying the gain of an external amplifier in the feedback path and without disturbing the core filter circuit. The shadow filter utilizes a CFCC block and three X+Copy ($X+_C$) CFCC blocks and its parameters can be independently tuned with the gain of the external amplifier. The gain of this external amplifier depends on the ratio of two grounded resistors. The workability of the proposed filter has been demonstrated in PSPICE using 180nm TSMC CMOS technology node.**

Keywords: Current mode circuits, Shadow filters, CFCC

I. INTRODUCTION

With advancements in technology, the demand for high-speed, power-efficient electronic circuits continues to rise [1]. In such a scenario, researchers are always on the lookout to develop and test new circuits that would serve the purpose. Since the development of current-mode circuits, they have continued to be the first choice of researchers, as they have many advantages over conventional voltage-mode circuits. Current-mode circuits offer better bandwidth performance, greater linearity, greater dynamic range, reduced circuitry, lesser chip area and lower power consumption. Another remarkable advantage that current mode circuits offer is that, summing operation can be realised simply by a circuit node, eliminating the need for a dedicated summing circuit, as in the case of voltage-mode circuits. One of the most popular current-mode active elements is the current conveyor block, first proposed by Sedra and Smith in 1968 [2]. The original current conveyor circuit has been modified since and used in a variety of situations. Cascading current follower and current conveyor blocks, we arrive at the Current Follower Current Conveyor (CFCC) block, a five terminal device that has current mirroring in its first and third stages and voltage mirroring at the second stage [1]. This property of the CFCC block allows us to design filter circuits for any required frequency response with the help of only grounded elements [3-5].

Shadow filter circuits find applications in the field of telecommunications wherein additional trimming is required for the filter parameters to adjust for the non-ideal effects without disturbing the original filter circuit. They are also used in radio frequency (RF) transmitters and receivers. They offer a great deal of flexibility as we can vary any one of the filter parameters orthogonally, by simply varying the gain of the external amplifier [6-8]. The external amplifier should be designed in such a manner that the gain depends only on the variations of grounded elements.

II. CFCC BLOCK

The input current at port P is copied to port Z with its direction reversed w.r.t. the input current. The potential present at port Z appears at the port I. The current present at port I is then copied into the output ports X+ and X- [3-5]. The current direction at X+ is the same as that of current at port I but has a reversed sense of direction at port X- with respect to current at port I. The Current Follower Current Conveyor (CFCC) block has been created by cascading current follower and current conveyor blocks as shown in Fig. 2. with the output of the current follower circuit being attached to the input of the current conveyor block [1]. The resulting block has both the output and input in current mode and can be used in many configurations owing to its flexible nature of voltage and current mirroring at its ports.

TABLE I. MOSFET W/L RATIOS FOR CFCC BLOCK WITH ADDITIONAL OUTPUT PORT $X+_C$.

MOSFET	W/L ratio
M1, M2, M29, M30	25/0.25
M3, M4, M31, M32	50/0.25
M5-M28	2.5/0.25
M33-M44	2.5/0.25

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:02:15 UTC from IEEE Xplore. Restrictions apply.

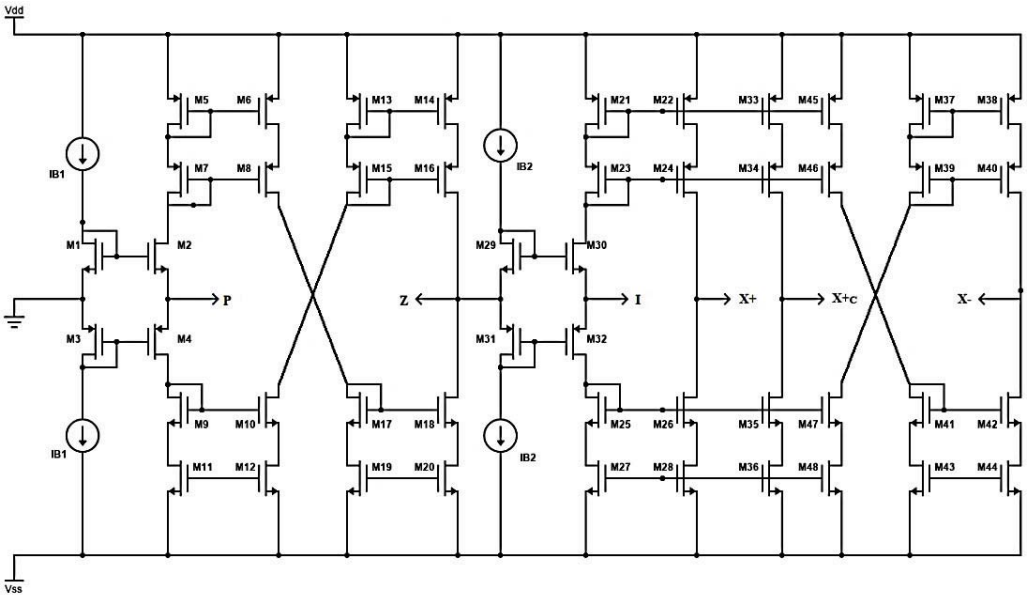


Fig. 1. MOS layout for CFCC block with additional output port $X+_C$.

Based upon the requirements, an additional output port $X+_C$ was added which generates a copy of the current at the X+ terminal. The introduction of this additional port helped in realizing the required filter using a minimum number of MOSFETs, thereby reducing chip area and power consumption.

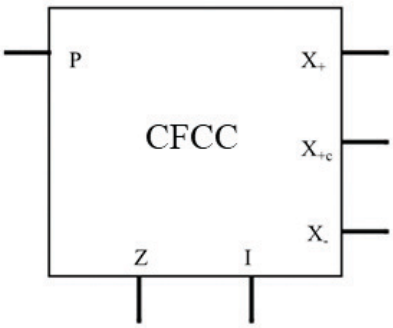


Fig. 2. CFCC Block with additional $X+_C$ port

The port relations for the standard CFCC block have been shown in the matrix:

$$\begin{bmatrix} V_p \\ I_z \\ V_i \\ I_{x+} \\ I_{x-} \\ I_{zc} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} I_p \\ V_z \\ I_i \\ V_{x+} \\ V_{x-} \\ V_{zc} \end{bmatrix} \qquad (I)$$

III. KHN FILTER USING CFCC

KHN filter can generate lowpass, high pass and bandpass output simultaneously from a single input [16]. It is a multi-feedback filter circuit. Using the CFCC with $X+_C$ port along with the conventional CFCC block, we have designed a KHN filter shown in Fig.3. The values of passive elements R1, R2, R3, R4, C1 and C2 have been calculated for a response with central frequency of 1MHz.

The relationship between output Currents I_{HP}. I_{LP}. I_{BP} and the input current I_{in} are given by the equations below:

$$I_{HP} = \left[\frac{As^2}{s^2 + \frac{A}{R_1C_1}s + \frac{A}{R_1R_2C_1C_2}} \right] I_{in} \qquad (1a)$$

$$I_{BP} = \left[\frac{\frac{A}{R_1C_1}s}{s^2 + \frac{A}{R_1C_1}s + \frac{A}{R_1R_2C_1C_2}} \right] I_{in} \qquad (1b)$$

$$I_{LP} = \left[\frac{\frac{A}{R_1R_2C_1C_2}}{s^2 + \frac{A}{R_1C_1}s + \frac{A}{R_1R_2C_1C_2}} \right] I_{in} \qquad (1c)$$

where $A = \frac{R_3}{R_4}$

IV. SHADOW FILTER USING CFCC

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:02:15 UTC from IEEE Xplore. Restrictions apply.

Introduced by Y. Lakys and A. Fabre, shadow filter is a second-order filter with independently electronically-tunable filter parameters [6]. In their paper, they demonstrated how the characteristic frequency of a second order filter could be independently tuned by introducing an external amplifier in the positive feedback path of the filter circuit [6-8].

By introducing a gain-tuneable amplifier in the feedback path as shown below in Fig. 4, we are able to create a filter whose center-frequency (f_0) and quality factor (Q) can be tuned electronically with the variation of only two grounded resistor elements [9-11]. The internal structure of the filter circuit remains the same and the bandwidth also remains unaffected throughout the operation [12-15].

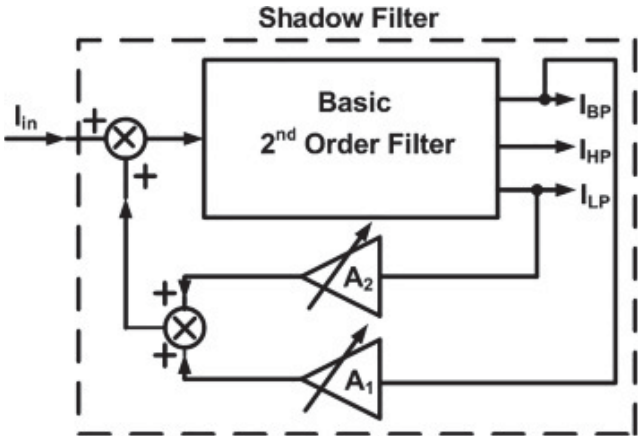


Fig. 3. Second Order Shadow Filter

Modified center-frequency for shadow filter,

$$\omega_0' = \sqrt{\frac{A(1-A')}{R_1R_2C_1C_2}} = \omega_0\sqrt{1-A'}$$

Modified quality-factor for shadow filter,

$$Q' = \frac{\omega_0\sqrt{1-A'}}{\frac{A}{R_1C_1}} = Q\sqrt{1-A'}$$

Bandwidth remains constant as before,

$$\Delta\omega' = \frac{\omega_0'}{Q'} = \frac{A}{R_1C_1} = \Delta\omega$$

Where $A = \frac{R_3}{R_4}$,

$$\omega_0 = \sqrt{\frac{A}{R_1R_2C_1C_2}},$$

and $Q = \sqrt{\frac{R_1C_1}{AR_2C_2}}$

The gain of the external amplifier in the feedback path dependent on the ratio of grounded resistances, $A' = \frac{R_5}{R_6}$.

V. SIMULATION RESULTS

The implementation of the shadow filter circuit in Fig. 3 has been corroborated with PSpice simulations. The results obtained for KHN filter with 1MHz of chosen center-frequency have been shown below in fig. 5(a). The results obtained for Shadow Filter by varying the feedback gain, A' have been also shown in Fig. 5(b) - 4(d).

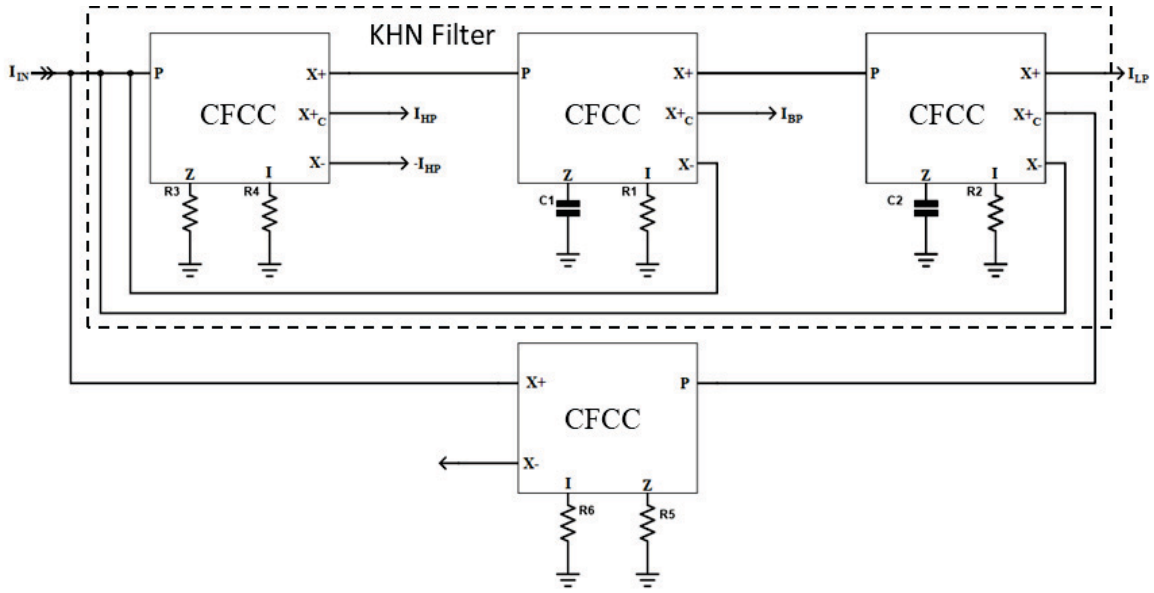


Fig. 1. Shadow Filter Using CFCC $X+_C$ and CFCC blocks

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:02:15 UTC from IEEE Xplore. Restrictions apply.

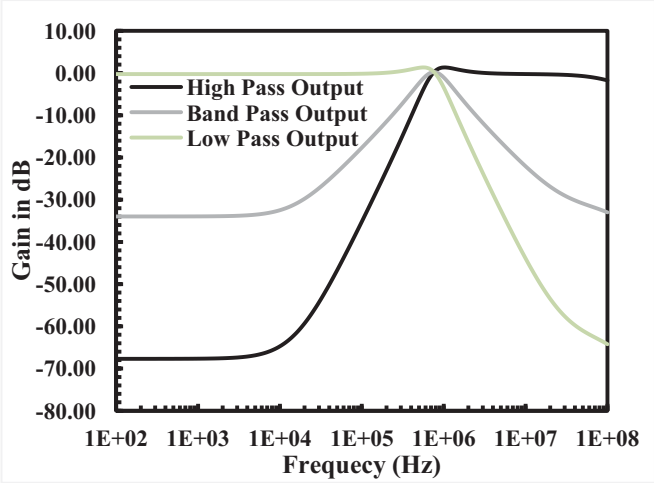


Fig. 5. (a) KHN Filter Response

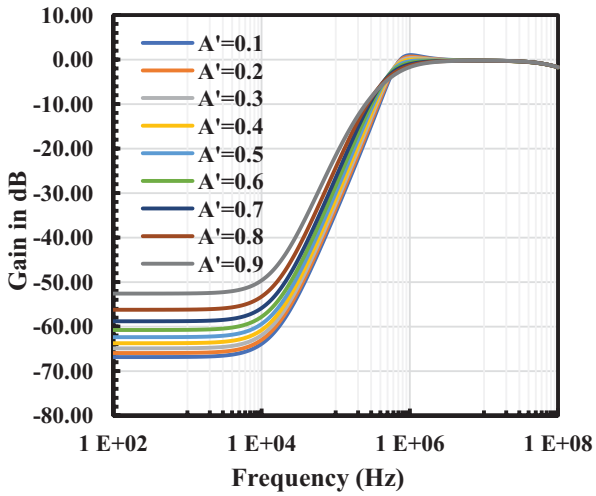


Fig. 5. (b) High Pass Shadow Filter Response

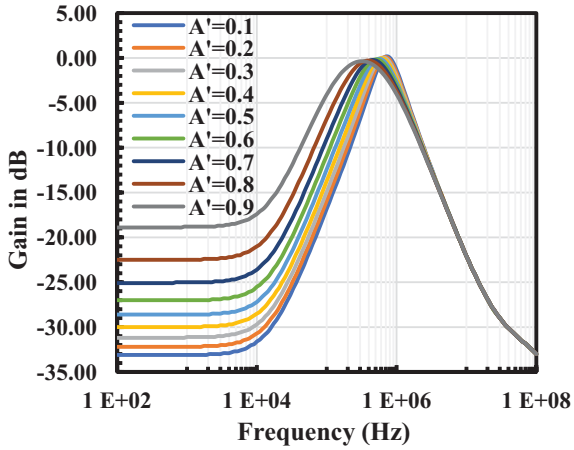


Fig. 5. (c) Band Pass Shadow Filter Response

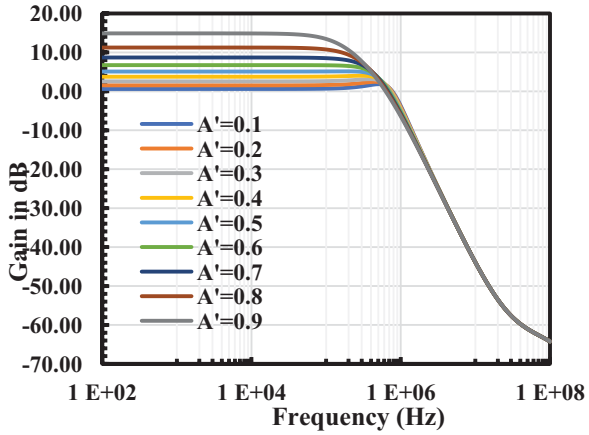


Fig. 5. (d) Low Pass Shadow Filter Response

VI. CONCLUSION

CFCC block was simulated in OrCAD PSpice software using 180nm technology model file. Shadow filter was then implemented using four CFCC blocks and grounded elements only. The gain of amplifier in the feedback path was made variable with the help of two grounded resistors and subsequent variation in center frequency and quality factor has been shown. The shadow filter was simulated for multiple values of gain of the external amplifier and the results obtained for the same have been incorporated.

REFERENCES

[1] D. Biolek, R. Senani, V. Biolkova and Z. Kolka, “Active Elements for Analog Signal Processing: Classification, Review, and New Proposals,” Radioengineering. 17, 2008.

[2] A. Sedra and K. Smith, “A second-generation current-conveyor and its applications,” IEEE Trans., vol. CT-17, pp. 132-134, 1970.

[3] A. K. Singh, P. Kumar and R. Senani, “Electronically tunable grounded/floating inductance simulators using Z-copy CFCCC,” Turkish Journal of Electrical Engineering & Computer Sciences. 26, pp. 1041-1055, 2018.

[4] A. K. Singh, P. Kumar and R. Senani, “New Grounded Immittance Simulators Employing A Single CFCC,” The Journal of Engineering, 2017.

[5] A. K. Singh, P. Kumar and R. Senani, “Current-mode Quadrature Oscillator Using CFCC,” American Journal of Electrical and Electronic Engineering, Vol. 6, No. 3, 2018, pp. 85-89.

[6] Y. Lakys and A. Fabre, “Shadow filters - new family of second-order filters,” Electronics Letters 46, 2010, pp. 276-277.

[7] Y. Lakys and A. Fabre, “Shadow filters generalization to nth-class,” Electron. Lett., 2010, 46, pp. 985–986.

[8] S. C. Roy, “Shadow Filters - A New Family of Electronically Tunable Filters,” IETE Journal of Education. 51, pp. 75-78, 2014.

[9] R. Anurag, R. Pandey, N. Pandey, M. Singh and M. Jain, "OTRA based shadow filters," 2015 Annual IEEE India Conference (INDICON), New Delhi, India, pp. 1-4, 2015.

[10] S. Buakaew, W. Narksarp and C. Wongtaychatham, "Fully Active and Minimal Shadow Bandpass Filter," 2018 International Conference on Engineering, Applied Sciences, and Technology (ICEAST), Phuket, Thailand, 2018, pp. 1-4.

[11] M. Abuelma'atti and N. Almutairi, “New CFOA-based shadow banpass filter,” pp. 1-3, 2016.

[12] F. Khateb, W. Jaikla, T. Kulej, M. Kumngern and D. Kubánek, “Shadow filters based on DDCC,” IET Circuits Devices Syst., 11, pp. 631-637, 2017.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:02:15 UTC from IEEE Xplore. Restrictions apply.

[13] P. Moonmuang, T. Pukkalanun, and W. Tangsrirat, “Voltage Differencing Gain Amplifier-Based Shadow Filter: A Comparison Study,” pp. 1-4, 2020.

[14] S. Buakaew, W. Narksarp and C. Wongtaychatham, "High Quality-Factor Shadow Bandpass Filters with Orthogonality to the Characteristic Frequency," 2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Phuket, Thailand, 2020, pp. 372-375.

[15] A. Upadhyay, "Shadow Filter for Orthogonal Modification of Center Frequency and Bandwidth with Constant Q Factor," 2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, Cambridge, UK, 2014, pp. 591-594.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:02:15 UTC from IEEE Xplore. Restrictions apply.

IOP Conference Series: Earth and Environmental Science

PAPER • **OPEN ACCESS**

Importance of Cytochrome P450 gene family from metabolite biosynthesis to stress tolerance: A review

To cite this article: N Laffaru Singpho and J G Sharma 2021 *IOP Conf. Ser.: Earth Environ. Sci.* **775** 012012

View the article online for updates and enhancements.

This content was downloaded from IP address 14.139.251.100 on 07/07/2021 at 07:36

Importance of Cytochrome P450 gene family from metabolite biosynthesis to stress tolerance: A review

N Laffaru Singpho, J G Sharma

Department of Biotechnology, Delhi Technological University, New Delhi 110042, India

namratalaffarusingpho_2k17/bt/15@dtu.ac.in

Abstract. CYP450 short for Cytochrome P450 is an enzyme superfamily involved in catalysing numerous biochemical reactions. It is known for its involvement in diverse plant processes. CYP450 is widely distributed in the case of eukaryotes and is said to exist in all domains of living organisms including bacteria, plants as well as mammals. The enzyme of this family plays a key role in the oxidative transformation of both endogenous as well as exogenous molecules. The growth and development of the plants are affected by several biotic and abiotic stresses regularly. They downgrade the crop quality and lead to a sharp decline in the productivity of the crop. CYP plays important role in providing protection to plants against these stresses. It does that by taking an active part in various detoxification as well as biosynthetic pathways. The objective of this review is to explore the role of CYP 450 in various metabolite biosynthesis as well as tolerance to various biotic and abiotic stresses in plants. This review aims to provide a framework for further investigation of the CYP450 gene family in plants and will also provide a strong base for the characterization of their diverse role in defence mechanisms against different abiotic and biotic stress and metabolite biosynthesis.

1. Introduction

Cytochrome P450 (CYP450) is a superfamily of enzymes commonly known for its key role in catalyzing various chemical reactions. Its traces have been seen in all domains of organisms. CYP450s are known to be widely distributed in the case of eukaryotes. It was Martin Klingenberg who first reported the CYP450 spectrum in the year 1958. The enzymes of this superfamily are responsible for the oxidative transformation of both exogenous as well as endogenous molecules [66], [23].

In previous studies, with the help of Genome-Wide Analysis, it was found that the Cytochrome P450 genes in the genomes of different organisms range from around 50 to 150. Besides It was also found that the plant *Arabidopsis thaliana* has 245 Cytochrome P450 [4], 57 Cytochrome P450 in the human genome, 285 Cytochrome P450 in the case of *Tartary Buckwheat* [96], and 174 Cytochrome P450 in case of *Morus notabilis* [55](Table 1). CYP450 is a diverse family with more than 35,000 members [66]. When CYP450 was at its initial stages of discovery it was considered to be a pigment rather than an enzyme. It was later found out that the particular pigment displayed vivid characteristics of a heme-protein [74]. Much later it was established that CYP450 were not pigments but enzymes. In CYP450, 450 nm is the wavelength of light the enzyme superfamily can strongly absorb and 'P' is short for pigment. CYPs can function in the carbon monoxide binding processes due to the presence of a particular heme group. This particular heme group is involved in the catalysis process in plants. In the case of plants, CYP450 was first identified in the cotton plant [25]. CYP450s are also known for

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

their important role in the metabolism of numerous pesticides and herbicides. Previous studies show that in the plant *Arabidopsis thaliana*, the P450 gene family vouches for being the 3rd largest gene family. In the case of plants, the molecular mass of the CYP450 gene family lies between 45 to 62kDa [102] with an average mass of approximately 55kDa [22]. Plant growth and development are challenged by numerous stresses daily. They can either be biotic or abiotic. These stresses affect the plants in many ways. They disturb the plant's physiological processes hence affecting its growth and development. They also downgrade the quality of crops which results in a sharp decline in crop yield. The CYPs are responsible for providing protection against these stresses by taking an active part in numerous biosynthetic as well as detoxification pathways [70]. When in danger, plants tend to produce some specific primary or secondary metabolites as a defense mechanism against various stress caused by insects, pathogens, temperature, weed, or even water. CYP helps in catalyzing the reaction during the biosynthesis of these metabolites and promotes the overall growth and development of plants.

Table 1 Distribution of CYP450s in different plants

Plants	**Total number of CYP450**	**Reference**
Triticum aestivum (Wheat)	1285	[50]
Gossypium hirsutum (Cotton)	672	[57]
Gossypium arboreum(Cotton)	379	-
Gossypium raimondii (Cotton)	374	-
Sorghum bicolor (Sorghum)	372	[68]
Oryza sativa (Rice)	356	-
Linum usitatissimum (Flaxseed)	334	[4]
Glycine max (Soybean)	332	-
Vitis vinifera (Grapevine)	315	[67]
Populus trichocarpa (Black cottonwood)	312	-
Fagopyrum tataricum (Tartary buckwheat)	285	[96]
Brassica oleracea (Cabbage)	279	[39]
Arabidopsis thaliana (Arabidopsis)	245	[4]
Solanum lycopersicum L. (Tomato)	233	[102]
Morus notabilis (Mulberry)	174	[55]

Nelumbo nucifera (Lotus)	172	[65]
Medicago truncatula (Barrel medic)	151	[49]

2. Role of CYP in plants

Apart from the role of CYPs in catalyzing various biochemical reactions mentioned earlier, CYP450s are also involved in numerous cellular processes which include growth and development of cell, primary and secondary metabolite biosynthesis as well as xenobiotics detoxification (Figure 1). They also facilitate the regulation of many important processes in a cell that eventually affects plant growth. CYP450s are responsible for the biosynthesis of compounds in numerous metabolic pathways. One example would be the involvement of the CYP706B1 gene in the isoprenoid pathway [53]. This particular hydroxylase is responsible for synthesizing gossypol. Gossypol is a toxic compound present in the cotton plant which makes it more resistant to insect stress due to its toxic nature. They are also responsible for the regulation of plant hormone metabolism. The hormone in turn manages the formation of leaves, stems, and flowers as well as covers the development and maturity of fruits. CYPs are involved in catalyzing secondary metabolites like steroids, flavonoids, glycoside, and many more [42]. These metabolites are in turn responsible for the overall development and growth of the plant. For example, structural genes in secondary metabolites such as flavonoids are often CYP450 genes [9]. Another key function of CYP450 is its important role in the biosynthesis of phytoalexins which is known for being a common plant defense response. Phytoalexins are released by plants during exposure to stress by parasites. They inhibit the growth and further development of parasites during a stressful environment.

3. Distribution of cyp450 in various plants

In plants, Cytochrome P450 is situated in the endoplasmic reticulum, mitochondria, and chloroplast [61]. CYP450 enzyme family is divided into two types i.e. a) A-Type- It consists of CYP71 clan and b) Non-A type- it consists of CYP746, CYP727, CYP711, CYP710, CYP97, CYP86, CYP85, CYP74, CYP72, CYP51[68], [6], [67] (Table 2). In plants, CYP is found in almost all its organs including the shoot, bulbs, hypocotyl, radicle as well as endosperm of germinating seeds. In addition to that in the case of both plants and animals, the spectral properties of CYPs are reported to be somewhat similar with the absorbance ranging from 447-452nm. Moreover, CYPs are also closely related to the biosynthetic pathway of the plant steroid hormone brassinolide (BR) [71].

There are four characteristics regions available that are common for all the P450s i.e.

a) I-helix
b) K-helix
c) Proline-rich membrane hinge and
d) “PERF” consensus

I-helix functions in process of oxygen-binding whereas “PERF” consensus ensures the stabilization of the core structure of the heme pocket [30].

In Figure 2 we can see the phylogenetic tree of the CYPP450 gene family in different species. The tree includes 6 *Arabidopsis thaliana*, 2 *Saccharomyces cerevisiae*, 2 *Solanum lycopersicum*, 2 *Brassica oleracea*, 2 *Oryza sativa*, and 1 *Zea mays* p450 genes. The analysis was done with the help of a computational biology tool named Mega 6.0 software. The nomenclature system states that the P450 belonging to the same family and same sub-family share around approximately 40% and 55% identity respectively.

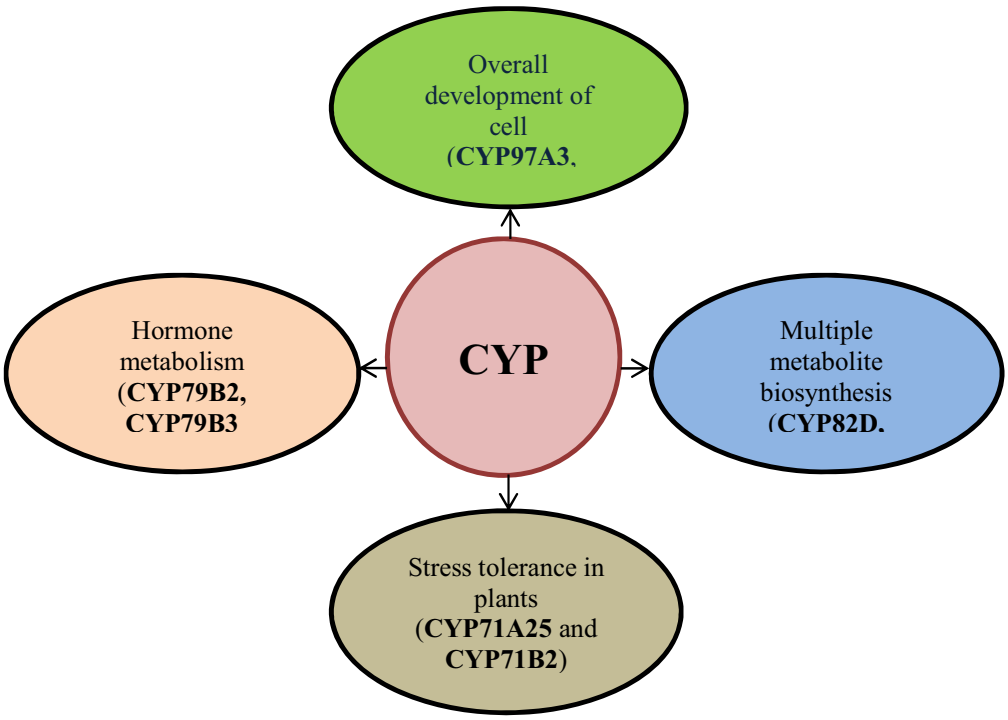


Figure 1 Different functions of CYP450 in plants

Table 2 Classification of CYP450 gene family in plants

Name of Clan	Name of Family	Name of Sub-family
CYP51	CYP51	CYPG, CYP51H
CYP71	CYP71	CYP71A, CYP71B, CYP71C, CYP71E, CYP71Z
	CYP73	CYP73A
	CYP75	CYP75A, CYP75B
	CYP76	CYP76A, CYP76B, CYP76M
	CYP77	CYP77A
	CYP78	CYP78A
	CYP79	CYP79A, CYP79B, CYP79D, CYP79F
	CYP80	CYP80A, CYP80G, CYP80B
	CYP81	CYP81A, CYP91B, CYP81D, CYP81F
	CYP82	CYPA, CYPC,
		CYPD, CYPG
	CYP83	CYP83A, CYP83B
	CYP84	CYP84A
	CYP89	NA
	CYP92	NA
	CYP93	CYP93A, CYP93B, CYP93C, CYP93E, CYP93G
	CYP98	CYP98A
	CYP99	CYP99A
	CYP701	CYP701A
	CYP703	CYP703A
	CYP705	CYP705A
	CYP706	CYP706B, CYP706M
	CYP712	NA

	CYP719	CYP719A, CYP719B
	CYP723	NA
	CYP726	NA
	CYP736	NA
CYP72	CYP72	CYP72A, CYP72B, CYP72C
	CYP709	CYP709B, CYP709C
	CYP714	CYP714A, CYP714B, CYP714D
	CYP715	NA
	CYP719	CYP719A, CYP719B
	CYP724	CYP724B
	CYP734	CYP734A
CYP74	CYP74	CYP74A, CYP74B, CYP74D
CYP85	CYP85	CYP85A
	CYP87	NA
	CYP88	CYP88A, CYP88D
	CYP90	CYP90A, CYP90B, CYP90C, CYP90D
	CYP702	NA
	CYP707	CYP707A
	CYP708	CYP708A
	CYP716	CYP716A
	CYP718	NA
	CYP720	CYP720B
	CYP724	CYP724B
	CYP725	CYP725A
	CYP728	NA
	CYP729	NA
	CYP733	NA
CYP86	CYP86	CYP86A, CYP86B
	CYP94	CYP94B, CYP94C
	CYP96	CYP96A, CYP96B
	CYP704	CYP704B
	CYP730	NA
	CYP731	NA
	CYP732	NA
CYP97	CYP97	CYP97A, CYP97B,CYP97C
CYP710	CYP710	CYP710A
CYP711	CYP711	CYP711A
CYP727	CYP727	NA

(The data given in the table is taken from [20], [42], [60], and [77])

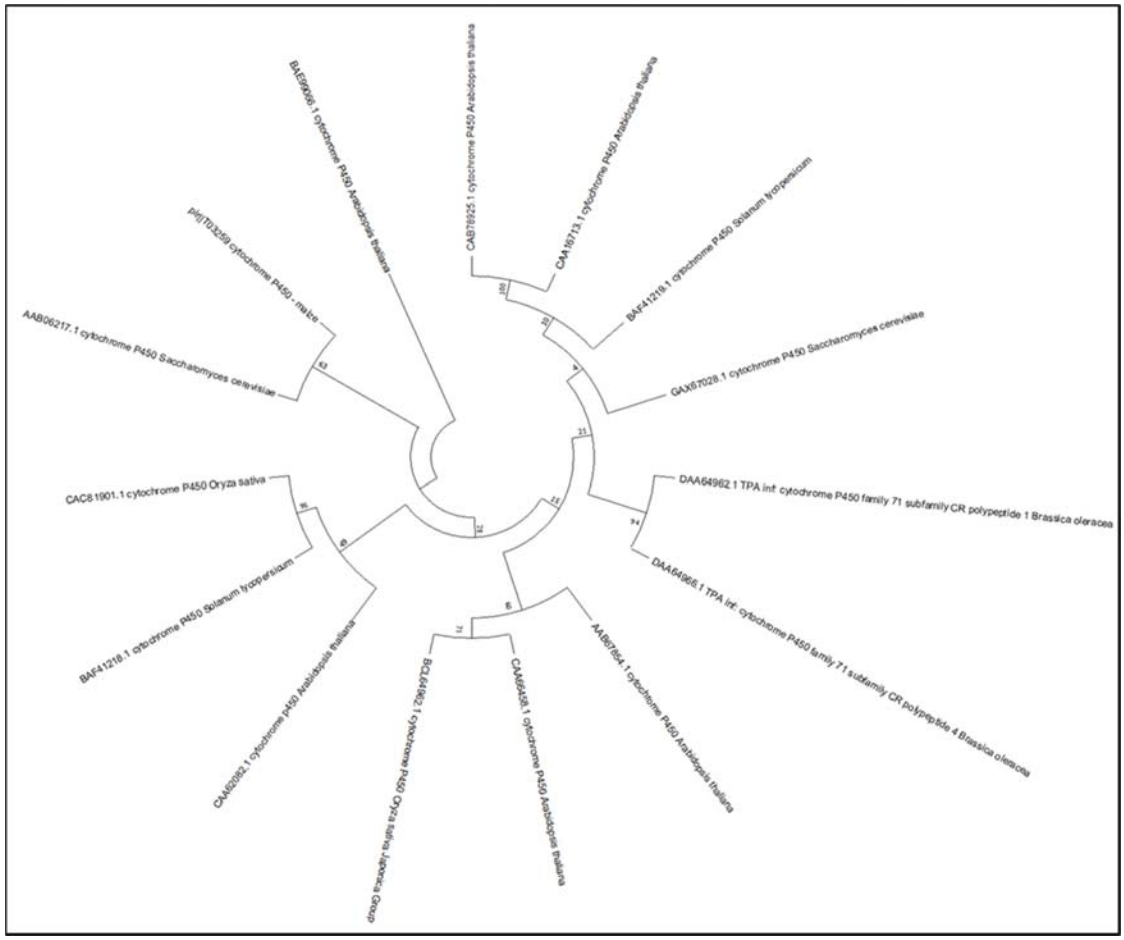


Figure 2 The phylogenetic tree of CYP450 in different species. The given tree includes 6 *Arabidopsis thaliana* P450 genes (CAB78925.1, CAA62082.1, CAA66458.1, CAA16713.1, AAB67854.1 and BAE99066.1), 2 *Saccharomyces cerevisiae* P450 genes (AAB06217.1 and GAX67028.1), 2 *Solanum lycopersicum* P450 genes(BAF41219.1 and BAF41218.1), 2 *Brassica oleracea* P450 genes(DAA64962.1 and DAA64966.1), 2 *Oryza sativa* P450 gene(BCL64962.1 and CAC81901.1) and 1 maize p450(PLRT03259) gene. It is constructed by MEGA 6.0 software using NJ algorithms with 1000 bootstrap replicates

4. Metabolite biosynthesis

Members of the CYP450 gene family are known for their diverse roles in core metabolisms. One of the key roles they play is their indirect participation in photosynthesis as well as photo-protection which is closely associated with the growth of plants [60]. Xanthophylls are yellow pigments that are responsible for light–harvesting during the photosynthesis process. CYP genes like CYP97A3 and CYP97C1 actively participates in the biosynthetic pathway of these yellow pigments as they function in catalyzing the process of hydroxylation of beta and gamma rings of carotenoids. CYP genes like CYP86A1, CYP86A2, CYP86A8, and CYP86A33 are responsible for playing a key role in the biosynthesis of cutin and suberin in plants like *Symphytum tuberosum* and Arabidopsis [108], [34], [93]. Cutin and suberin are complex biopolymers that are known for protecting plants from certain disturbances including pathogens and UV as they are capable of controlling processes such as water evaporation as well as water transpiration. Cutin majorly consists of glycerol and fatty acids. Suberin on the other hand consists of fatty acid-derived domains. Now, the CYP86 gene family helps in the biosynthesis of these biopolymers as they function in the hydroxylation process of fatty acids which is essentially the most vital component of these compounds.

In plants, members of the CYP450 gene family like CYP93B, CYP93E, and CYP93G are responsible for the secondary metabolites biosynthesis in plants [20]. Additionally, the CYP450 gene family is also considered to be an essential component of auxin biosynthesis in certain plants [5]. Auxins are growth-inducing hormones that function in certain tropic behaviors, root initiation, and also apical dominance [22]. CYP734, CYP724, and CYP90 are involved in the process of glycoalkaloids biosynthesis. The glycoalkaloids are usually known to be toxic to living beings however they are an essential part of the Solanaceae family. In potato plants, CYP74D genes play a vital role in the biosynthesis of antimicrobial compounds such as oxylipins [95]. These antimicrobial compounds are derivatives of fatty acids and are considered useful for overall plant growth. Oxylipins also provides defense against certain stresses in the case of plants. Certain CYP genes like AtCYP97C1 and AtCYP97A3 are responsible for determining the pathway of lutein in the Sorghum plant [43]. Lutein in plants functions in the protection of chlorophyll due to its participation in the absorption and transfer of energy. (Table 3)

Table 3 Different CYP genes functioning in metabolites biosynthesis

Plant Species	**CYP genes**	**Function of CYP450**	**Reference**
Arabidopsis thaliana	AtCYP72B1, AtCYP72C1	Modulates signal transduction in this particular plant species	[100]
-	CYP79B2, CYP79B3, CYP83	Biosynthesis of the plant hormone auxin	[5]
-	CYP88A, CYP714A	Gibberellin metabolic pathways	[118]
-	CYP703A2	Development of pollen	[62]
-	CYP707A	Modulates plant hormone abscisic acid	[59]
-	AtCYP74A	Jasmonic acid biosynthesis	[78]
-	CYP72A31	Provides defense against herbicides	[87]
Oryza sativa	CYP72A31	Provides defense against herbicides	-
Oryza sativa	CYP93B, CYP93E, CYP93G	Flavonoid biosynthesis	[20]
Hordeum vulgare	CYP707A	Modulation of plant hormone abscisic acid	[59]

Sorghum bicolor	CYP79	Cyanogenic glucosides biosynthesis	[7]
Sorghum bicolor	AtCYP97C1, AtCYP97A3	Determines the pathway for biosynthesis of lutein	[43]
Ocimum basilicum	CYP82D	Secondary metabolites biosynthesis	[9]
Nicotiana rustica	CYP51	Triterpenes synthesis	[26]
Vitis vinifera	CYP75	Expression of flavonoids in plants	[12]
Solanum lycopersicum	CYP90B3, CYP724B2	Plant steroid hormone BR (Brassinolide) biosynthetic pathway	[71]
Solanum lycopersicum	AtCYP90A1, AtCYP90B1	Increases vegetative growth along with seed production	[14]
Solanum tuberosum	CYP74D	Synthesis of antimicrobial compound namely oxylipin	[95]
Hordeum vulgare	CYP71A11, CYP71AH11, CYP73A1, CYP81B1, CYP81B2, CYP76B1	Chlortoluron metabolism	[70]
-	CYP96B22	Biosynthesis of wax	[17]
Gossypium arboreum	CYP706B1	Modification of isoprenoid pathway	[53]
Symphytum tuberosum	CYP86A33	Suberin biosynthesis	[93]
Legumes	CYP93C	Legume specific isoflavonoid biosynthetic pathway	[20]
Land plants	CYP93B, CYP93E, CYP93G	Saponin biosynthesis	[63]

5. ROLE OF CYPs IN DIFFERENT TYPES OF STRESS TOLERANCE IN PLANTS

5.1. Abiotic stress

The stress that facilitates the damaging of plants by the action of certain environmental factors such as availability of water, fluctuations in the salt concentration as well as temperature, increasing metal toxicity, etc. is called Abiotic stress. In plants, it is responsible for the loss of crop yield and decrease in overall productivity. CYP is responsible for protecting plants against various stresses caused by environmental factors (Table 4).

5.1.1. Drought stress. Water or drought stress is the lack of moisture that is considered to be a minimum requirement for the normal growth of an average plant. The main factors responsible for causing drought stress include low rainfall, frequent fluctuation of temperature from low to high or vice-versa, salinity, and many others. Drought stress indirectly affects plant productivity. It leads to reduced cell division in plants. It also disturbs the plant's stomatal oscillation along with its nutrient and water consumption process. It is known to be one of the most common stresses. The survivability of plants during drought stress depends on the maintenance of cell homeostasis in a water-deficient environment. When under such stress, plants trigger certain hormonal activity as a resistance mechanism. One example of such plant hormone is ABA. ABA plant hormone is responsible for activating multiple stress-responsive genes. During drought stress, this particular hormone gets isomerized into phaseic acid. This process is catalyzed by CYP enzyme CYP707A [119]. In certain plants, CYP genes are responsible for the up-regulation of the water inlet in a water-deficit environment. During drought stress, CYP synthesizes leaf lignin which forms the key structural materials that support tissues of plants and is also responsible for the grain formation [35]. One example of the CYP450 gene family which is involved in the defense mechanism against drought stress is CYP96A8. It is responsible for the biosynthesis of lignin in few plants. In the plant *Arabidopsis thaliana*, some CYP genes like CYP86A2 showed properties such as reduction of cuticle membrane thickness. This particular property increases the tolerance of plants against stress caused by water deficiency by increasing the water permeability (Xiao et al., 2004).

5.1.2. Salinity stress. Salinity or salt stress is the stress caused due to the abundance of salt in a plant habitat. Similar to drought stress, it also affects the overall plant productivity by affecting the germination of a plant which eventually ends up affecting the crop yield. It is one of the abiotic stresses which hinders the smooth development of plant all over the globe. The major factors causing salinity stress include the use of poor quality water for the purpose of salinization as well as irrigation. High salinity gives rise to several other stresses such as oxidative and water stress. It also leads to nutritional disorders as well as ion toxicity. Other disadvantages of salinity stress include genotoxicity and reduced cell expansion and division. CYP indirectly provides protection to plants against stress caused due to fluctuation in salt concentration by taking an active part in the ABA biosynthesis as the lower is the ABA level; the higher becomes the tolerance against stress caused by salt [38]. Previous studies state that alteration in CYP expression can lead to tolerance against salt stress in the case of plants. Reactive oxygen species (ROS) homeostasis and leaf Na+ exclusion are the two most widely known mechanisms related to this particular stress [115]. CYPs are also known for the regulation of ROS homeostasis resulting in providing salinity tolerance to the plants. For example, in the maize plant, the TaCYP81D5 gene increases the plant's tolerance to this particular stress by speeding up the activity of ROS scavenging and is also involved in the seedling activity [104].

5.1.3. Heavy metal toxicity. Heavy metal toxicity refers to the oxidative stress plants undergo when they are exposed to heavy metals. Metals have harmful biological effects since they are non-decomposable in nature. Heavy metal includes metals such as Nickel, Zinc, Cobalt, Copper, Thallium, Arsenic, Silver, and others. Heavy metal toxicity damages the cell assemblies of the plant. Among all the heavy metals, five are considered to be most lethal and are treated with utmost care as they have

the highest toxicity. Those five heavy metals are Lead, Mercury, Chromium, Arsenic, and Cadmium [21]. Improper disposal of sewage sludge is one of the causes of heavy metal toxicity [3]. CYPs are responsible for the metabolism of these heavy metals. Due to enhanced metabolism, few CYP genes such as CYP2E1 are capable of withstanding mercury toxicity in certain plants like *Medicago sativa*. In addition to that, they also function in the secondary metabolites biosynthesis such as alkaloids, flavonoids, etc. [91]. In the maize plant, the CYP88A gene facilitates gibberellin synthesis [32]. Gibberellins are growth hormones that function in the integration of numerous hormone signaling pathways when exposed to heavy metal toxicity stress. A high level of metal toxicity also disturbs the process of photosynthesis in plants. When exposed to heavy metal stress, P450s actively participates in the process of detoxification [97].

5.1.4. Temperature stress. Temperature stress is the increase or decrease in temperature which affects a plant's biochemical as well as physiological processes. It can either be cold stress (less than 20 degree C) or heat stress (approx. 15 degree C above the optimum temperature). Temperature stress also includes freezing stress i.e. the extreme form of cold stress. Here the temperature is below 0 degree C [89]. Since plants maintain themselves at a constant optimum temperature, change in temperature even for the slightest duration may affect them greatly. Temperature stress in plants leads to certain mishaps such as reduction of stability of cell membrane, poor germination, protein denaturation, rapid unfolding of nucleic acid, yellowing of leaf, and loss of enzyme activity [117]. Heat stress causes protein denaturation by leading to the breakdown of cell proteins in plants. It also affects the permeability of the membrane as extreme heat stress results in the melting of cell walls in the case of plants [101]. They eventually lead to loss of crop yield and a sharp decline in crop productivity. Cold stress in plants results in dehydration and starvation as cold stress affects the rate of uptake of essential nutrients and most importantly water. Various studies have indicated that CYP genes up-regulate under temperature stress. For example, In *Lolium perenne*, it was found that CYP73A, CYP75B, CYP75A genes were up-regulated due to fluctuations in the temperature [99] and In sorghum up-regulation of CYP99A1 and CYP709C1 genes under temperature stress was seen [15]. In *Brassica napus*, CYP71A23 is responsible for pollen sterility under temperature stress [83]. In the Arabidopsis plant, the CYP83A1 gene shows a 2-4 fold expression under temperature stress [10].

Table 4 Diverse functions of CYP450 in different plants species when exposed to various abiotic stresses

Type of stress	Plant	CYP genes	Role of CYP under various abiotic stress	Reference
Drought stress	*Sorghum bicolor*	CYP71A25, CYP71B2	CYP up-regulates under drought stress	[41]
	Zea mays	CYP96A8	Biosynthesis of Lignin	[35]
	Arabidopsis thaliana	CYP707A1, CYP707A2	CYP707A1 and CYP707A2 gene is up-regulated under water-deficient condition	[47]
	-	CsCYT75B1	Enhances antioxidant activity	[85]
	-	CYP707A	ABA synthesis	[119]

	-	CYP86A2	Reduction of cuticle membrane thickness	[108]
	Citrus sinensis	CYP75B1	CsCYP75B1 gene up-regulates under drought stress	-
Salinity stress	*Zea mays*	TaCYP81D5	CYP speeds up the ROS scavenging activity	[104]
	Arabidopsis thaliana	CYP709B3	AtCYP709B3 exhibit increased expression under salinity stress	[38]
	Physcomitrella patens	NA	Reduces the damaging of cell wall under high salinity stress	[106]
	Robinia pseudoacacia	CYP709	CYP genes showed increased expression when exposed to salinity stress	[110]
Heavy metal toxicity	*Panax ginseng*	CYP71	Biosynthesis of secondary metabolites	[91]
	Zea mays	CYP88A	Gibberellin biosynthesis	[32]
	Medicago sativa		Phytoremediation	[116]
	Arabidopsis thaliana	CYP83A1	CYP exhibits 4-fold expression against Al toxicity	[29]
	Triticum aestivum	CYP88A	Development of shoot	[32]
	Oryza sativa	CYP714A3	Development of shoot	[32]
	Pisum sativum	CYP88A	CYP88A gene showed stunted growth under heavy metal toxicity	[113]

Temperature stress	*Panicum maximum*	CYP71A23	When under heat stress CYP genes up-regulates	[107]
	Lolium perenne	CYP73A, CYP75B, CYP75A	CYP73A, CYP75B, CYP75A genes are up-regulated under temperature stress	[99]
	Arabidopsis thaliana	CYP83A1	CYP83A1 gene exhibits 2-4 fold expression under heat and cold stress	[10]
	Brassica napus	CYP71A23	Pollen sterility	[83]
	Sorghum bicolor	CYP99A1, CYP709C1	Up-regulation of CYP99A1 and CYP709C1 genes under temperature stress	[15]

5.2. Biotic stress

The stress that results in damaging of plants by the actions of living organisms including insects, bacteria, weeds, etc. which disturbs the overall dynamic of plants is called biotic stress. Biotic stresses in plants usually refer to hindrances such as insects, diseases, weeds, etc. These stresses cause deprivation of nutrients and minerals required by the host plant which in some extreme cases also lead to plant death. Apart from disturbing plant growth, biotic stresses are also responsible for affecting the seed quality, rotting of root, and downgrading the crop yield. In plants such as *Zea mays* and *Pisum sativum*, CYP genes were found to be up-regulated during the time of wounding [79], [24]. Additionally, the CYP gene i.e. CYP96A15are also closely involved with a very common structural plant defence mechanism called epicuticular wax biosynthesis. (Table 5)

5.2.1. Disease stress. Disease stress is considered the most prominent out of all these biotic stresses. Fungi are major contributors of disease stress in plants. It is closely followed by plant stress caused by insects. In the Arabidopsis plant, a member of the CYP450 gene family AtCYP76C2 gene results in rapid cell death as a coping response to bacterial stress caused by *Pseudomonas syringae* as it restricts the further spreading and growth of pathogens [28]. In *Capsicum annuum*, caCYP1 is a key player as it is known for its involvement in hypersensitive response (HR) to protect the plant from infection by *Xanthomonas axonopodis* [44]. It is also responsible for the biosynthesis of cutin as well as cyanogenic glucosides [108]. One of the many factions of CYP in plants includes the involvement of CYP93C genes in providing protection against disease stress. This particular gene enhances the intake of isoflavonoids by plants thus facilitating the tolerance against disease stress. CYP450 genes like CYP81E1, CYP81E3, CYP81E7, and CYP93A1 are also considered key players in the isoflavonoids biosynthesis in the case of plants like *Medicago truncatula, Glycyrrhiza echinata, Glycine max L* and *Cicer arietnum L* [51], [2], [90], [75]. Isoflavonoids in turn provide protection against diseases by increasing dietary intake. In the Arabidopsis plant, CYP71B15 is responsible for the biosynthesis of camalexin and by doing so ensures the protection of plants from disease stress [22] as the secondary

metabolite camalexin discourages the further development of bacterial and fungal pathogens. Camalexins are known for their defence mechanism against a virulent bacterial pathogen namely Pseudomonas syringae. Another example of the involvement of CYP as a defence against pathogens in Arabidopsis is that this plant usually produces phytoalexins as a response to attacks by bacterial pathogens. CYP genes like CYP71B15, CYP71A12, CYP79B3, CYP79B2, and CYP71A13 in the said plant are responsible for the biosynthesis of the previously mentioned aromatic compound camalexin [22]. Hence, helps in the defence mechanism against disease stress.

5.2.2. Insect stress. Insect stress causes physical damage to plants. They also serve as carriers of numerous bacteria and viruses. Insect stress in plants also disturbs important processes such as photosynthesis as insects tend to reduce the surface area of a leaf which is directly related to the rate of photosynthesis. They function in the biosynthesis of xanthophyll. This pigment functions in the light-capturing process during photosynthesis. CYP is responsible for increasing the would-induced defence during the biotic stress environment. In the sorghum plant, CYP71E1 and CYP79A1 catalyse the reaction of conversion on tyrosine, which in turn optimizes pest resistance as well as food safety [5]. Additionally, in few plants like Arabidopsis, CYP450s like CYP86A1 are also known for their function in the biosynthesis of suberin which is a defence mechanism against insect stress [34]. Arabidopsis plants also include CYP PAD3 genes which make the plant resistant to *Myzus persicae* or green peach aphid [81]. The green peach aphids are considered a major threat to plants all around the world due to their ability to work as a vector carrying different plant viruses. Another important role of CYP450 is their involvement in Berberine biosynthesis. Berberines are secondary metabolites that function in providing protection to plants from Plutella xylostella commonly known as the Diamond back moth. This insect is known for its destructive nature and high tolerance to bio-pesticides as well as certain chemicals. Berberine is an alkaloid that is known to be toxic to insects including the diamond back moth. CYP719 genes are responsible for the biosynthesis of this alkaloid in *Coptis japonica* [36]. Hence, they provide defence against insect stress in plants.

5.2.3. Weed stress. Another type of biotic stress includes weeds or undesirable and unprofitable plants. Weed is one of the biotic stresses that the plant faces which yet again challenges its productivity, growth as well as development as there is clear competition for water, minerals, nutrients as well as sunlight between weed and the plant. Weed is a potential threat to crop productivity as they dominate the environment because of their rapid growth criterion. Certain plants such as Sorghum release allelochemicals to inhibit weed growth. CYP71AM1 gene synthesizes these allelochemicals in such plants. It is responsible for catalysing sorgoleone which is an allelochemicals [76]. CYP71C is considered to be vital in the process of biosynthesis of benzoxazinoids namely DIBOA and DIMBOA which are specialized metabolites responsible for inhibiting the weed growth in the case of the wheat plant [69]. Apart from their role in providing defence against weed they also protect the plants from other biotic stresses such as insects and diseases.

Table 5 Diverse role of CYP in different plants under various biotic stresses

Stress	**Plant**	**CYP genes**	**Role of CYP under various biotic stress**	**Reference**
Disease stress	*Arabidopsis thaliana*	AtCYP76C2	Results in the rapid cell death as a coping response to bacterial stress caused by *Pseudomonas syringae*	[28]

	Arabidopsis thaliana	CYP71B15	Biosynthesis of secondary metabolite camalexin which in turn provide protection against bacterial pathogen *Pseudomonas syringae*	[22]
	-	CYP71B15, CYP71A12, CYP79B3, CYP79B2, CYP71A13	Modulates the level of phytoalexins resulting in tolerance against some pathogen	[22]
	Nicotiana benthamiana	GmCYP82A3	High tolerance to gray mold and black shank	[111]
	Capsicum annuum	CaCYP1	Involvement in hypersensitive response (HR) to protect plant from infection by *Xanthomonas axonopodis*	[44]
	Triticum aestivum	CYP709C3v2	Tolerance to Fusarium head blight	[46]
	Cicer arietnum L.	CYP81E3	Biosynthesis of isoflavonoid. Isoflavonoid in turn enhances the dietary intake hence provides protection from diseases.	[75]
	Glycine max L.	CYP93A1	-	[90]
	Glycyrrhiza echinata	CYP81E1	-	[22]
Insect stress	*Arabidopsis thaliana*	CYP86A1	Suberin biosynthesis	[34]
	-	CYP86A2, CYP86A8	Cutin biosynthesis	[108]
	-	CYP PAD3	Protection from *Myzus persicae* commonly known as green peach aphid in the case of Arabidopsis plant	[81]
	Nicotiana tobacum	Uncharacterized	Resistance to aphids	[103]

	Oryza sativa	CYP704B2	Anther cutin biosynthesis	[48]
	Symphytum tuberosum	CYP77A20, CYP77A19	Biosynthesis of suberin	[93]
	Sorghum bicolor	CYP71E1, CYP79A1	Conversion of tyrosine during the biosynthesis of Dhurrin to ensure resistance against pest	[7]
	Coptis japonica	CYP719	CYP functions in Berberine biosynthesis to ensure resistance against Diamond black moth	[36]
	Conifers	CYP720B4	Biosynthesis of resin acid	[33]
Weed stress	*Sorghum bicolor*	CYP71AM1	Biosynthesis of allelochemicals to suppress the growth of weeds	[76]
	Triticum aestivum	CYP71C	Biosynthesis of DIBOA, DIMBOA to inhibit further growth of weeds	[69]
Other biotic stresses	*Glycine max*	CYP82A3	Play important role in the signaling pathways	[111]
	Zea mays	NA	Up-regulation of CYP genes during wounding	[79]
	Pisum sativum	NA	CYP genes get up-regulated during wounding	[24]
	Populus trichocarpa	CYP79D	Aldoximes biosynthesis	[37]

6. Analysis/ Discussion

CYPs are considered vital in a plant's defence mechanism against different stresses and help in catalysing the reaction during the biosynthesis of primary and secondary metabolites. A CYP first figure out the external stress condition and generates cellular responses accordingly. CYP450 play important role in the biosynthesis of phytoalexins which are released by plants during exposure to stress by parasites. Moreover, CYPs are also involved in important cellular processes like photosynthesis. Xanthophylls are yellow pigments that are responsible for light–harvesting during the photosynthesis process. CYP genes like CYP97A3 and CYP97C1 actively participates in the

biosynthetic pathway of these yellow pigments as they function in catalyzing the process of hydroxylation of beta and gamma rings of carotenoids. In the plant *Arabidopsis thaliana*, CYP86A2 showed properties such as reduction of cuticle membrane thickness which results in the increased tolerance of plants against stress caused by water deficiency by increasing the water permeability. Gibberellins are growth hormones that function in the integration of numerous hormone signaling pathways when exposed to heavy metal toxicity stress. In plants such as maize CYPs facilitates synthesis of these growth hormones. Another important role of CYP450 is their involvement in Berberine biosynthesis. Berberines are secondary metabolites that function in providing protection to plants from *Plutella xylostella* commonly known as the Diamond back moth. This insect is known for its destructive nature and high tolerance to bio-pesticides as well as certain chemicals. Berberine is an alkaloid that is known to be toxic to insects including the diamond back moth.

7. Conclusion

Environmental stresses caused via biotic and abiotic factors often hinder the growth as well as effects the development of plants. Living organisms comprising fungi, bacteria, insects, etc. are often the cause behind biotic stresses. Abiotic and biotic stresses such as these occur due to a mixture of physical as well as environmental factors widely affecting the plants by causing stress thereby reducing plant productivity. However, plants have developed many defense mechanisms against these stresses. In plants, the Cytochrome P450 gene family is responsible for providing protection to plants against multiple abiotic and biotic stresses by actively taking part in numerous detoxification and biosynthetic pathways. They are also responsible for the biosynthesis of numerous compounds to ensure the survival of the plant when exposed to certain stresses. It is expected that factors causing abiotic and biotic stress will increase in the coming future. One example is the rapid increase in Earth's temperature. It would become a necessity to breed plants that are resistant to these stresses. This study facilitates the characterization of CYP450 in plants to observe its role in metabolite biosynthesis and defense against different stresses in plants. Furthermore, this study will also provide us with a solid foundation for further analysis of the CYP450 gene in numerous plant species in the future. With an increasing amount of genomic information, our knowledge of this specific enzyme will keep on increasing as well. We have previously discussed the diverse functions of CYP450 in plants. Since one of its many functions is its role in generating defense response under stress conditions, increasing our knowledge of this enzyme superfamily will help us develop plants with magnified tolerance to such stresses in the coming future.

Acknowledgement

The author wishes to thank Dr. Nirala Ramchiary, Dr Jai Gopal Sharma and the Department of Biotechnology, Delhi Technological University for their kind guidance.

References

[1] Ai J, Zhu Y, Duan J, Yu Q, Zhang G, Wan F and Xiang Z H 2011 Genome-wide analysis cytochrome P450 monooxygenase genes in the silkworm, Bombyx mori *Gene* **480(1-2)** 42-50

[2] Akashi T, Aoki T and Ayabe S I 1998 CYP81E1, a cytochrome P450 cDNA of licorice (Glycyrrhiza echinataL.), encodes Isoflavone 2′-hydroxylase *Biochemical and biophysical research communications* **251(1)** 67-70

[3] Ali H, Khan E and Sajad M A 2013 Phytoremediation of heavy metals—concepts and applications *Chemosphere* **91(7)** 869-81

[4] Babu P R, Rao K V and Reddy V D 2013 Structural organization and classification of cytochrome P450 genes in flax (Linum usitatissimum L.) *Gene* **513(1)** 156-62

[5] Bak S and Feyereisen R 2001 The involvement of two P450 enzymes, CYP83B1 and CYP83A1, in auxin homeostasis and glucosinolate biosynthesis *Plant Physiol* **127(1)** 108-18

[6] Bak S, Beisson F, Bishop G, Hamberger B, Höfer R, Paquette S and Werck-Reichhart D 2011 Cytochromes P450 *The Arabidopsis Book/American Society of Plant Biologists* **9**

[7] Bak S, Kahn R A, Nielsen H L, Møller B L and Halkier B A 1998 Cloning of three A-type cytochromes P450, CYP71E1, CYP98, and CYP99 from Sorghum bicolor (L.) Moench by a PCR approach and identification by expression in Escherichia coli of CYP71E1 as a multifunctional cytochrome P450 in the biosynthesis of the cyanogenic glucoside Dhurrin *Plant molecular biology* **36(3)** 393-405
[8] BancosİS, Nomura T, Sato T, Molnár G, Bishop G J, Koncz C and Szekeres M 2002 Regulation of transcript levels of the Arabidopsis cytochrome P450 genes involved in brassinosteroid biosynthesis *Plant physiol* **130(1)** 504-13
[9] Berim A and Gang D R 2013 The roles of a flavone-6-hydroxylase and 7-O-demethylation in the flavone biosynthetic network of sweet basil *Journal of Biological Chemistry* **288(3)** 1795-805
[10] Bilodeau P, Udvardi M K, Peacock W J and Dennis E S 1999 A prolonged cold treatment-induced cytochrome P450 gene from Arabidopsis thaliana Plant, Cell & Environment 22(7) 791-800
[11] Bolwell G P, Bozak K and Zimmerlin A 1994 Plant cytochrome P450 *Phytochemistry* **37(6)** 1491-506
[12] Castellarin S D, Di Gaspero G, Marconi R, Nonis A, Peterlunger E, Paillard S, Testolin R 2006 Colour variation in red grapevines (Vitis vinifera L.): genomic organisation, expression of flavonoid 3'-hydroxylase, flavonoid 3', 5'-hydroxylase genes and related metabolite profiling of red cyanidin-/blue delphinidin-based anthocyanins in berry skin *BMC Genomics* **7(1)** 1-17
[13] Chen J, Song Y, Zhang H and Zhang D 2013 Genome-wide analysis of gene expression in response to drought stress in Populus simonii *Plant molecular biology reporter* **31(4)** 946-62
[14] Choe S, Fujioka S, Noguchi T, Takatsuto S, Yoshida S and Feldmann K A 2001 Overexpression of DWARF4 in the brassinosteroid biosynthetic pathway results in increased vegetative growth and seed yield in Arabidopsis *The Plant Journal* **26(6)** 573-82
[15] Chopra R, Burow G, Hayes C, Emendack Y, Xin Z and Burke J 2015 Transcriptome profiling and validation of gene based single nucleotide polymorphisms (SNPs) in sorghum genotypes with contrasting responses to cold stress BMC genomics **16(1)** 1-11
[16] Coon M J 2005 Cytochrome P450: nature's most versatile biological catalyst *Annu. Rev. Pharmacol. Toxicol* **45** 1-25
[17] Delventhal R, Falter C, Strugala R, Zellerhoff N and Schaffrath U 2014 Ectoparasitic growth of Magnaporthe on barley triggers expression of the putative barley wax biosynthesis gene CYP96B22 which is involved in penetration resistance *BMC plant biol* **14(1)** 1-14
[18] Du H, Huang Y and Tang Y 2010 Genetic and metabolic engineering of isoflavonoid biosynthesis *Appl. Microbiol. Biotech* 86(5) 1293-312
[19] Du H, Ran F, Dong H L, Wen J, Li J N and Liang Z 2016 Genome-wide analysis, classification, evolution, and expression analysis of the cytochrome P450 93 family in land plants *PloS one* **11(10)** e0165020
[20] Du Y, Chu H, Chu I K and Lo C 2010 CYP93G2 is a flavanone 2-hydroxylase required for C-glycosylflavone biosynthesis in rice *Plant physiol* **154(1)** 324-33
[21] Dutta S, Mitra M, Agarwal P, Mahapatra K, De S, Sett U and Roy S 2018 Oxidative and genotoxic damages in plants in response to heavy metal stress and maintenance of genome stability *Plant signaling & behavior* **13(8)** e1460048
[22] Feldmann K A 2001 Cytochrome P450s as genes for crop improvement *Current opinion in plant biology* **4(2)** 162-67
[23] Fischer M, Knoll M, Sirim D, Wagner F, Funke S and Pleiss J 2007 The Cytochrome P450 Engineering Database: a navigation and prediction tool for the cytochrome P450 protein family *Bioinformatics* **23(15)** 2015-17
[24] Frank M R, Deyneka J M and Schuler M A 1996 Cloning of wound-induced cytochrome P450 monooxygenases expressed in pea *Plant physiol* **110(3)** 1035-46
[25] Frear D S, Swanson H R and Tanaka F S 1969 N-demethylation of substituted 3-(phenyl)-1-methylureas: isolation and characterization of a microsomal mixed function oxidase from cotton *Phytochemistry* **8(11)** 2157-69
[26] Geisler K, Hughes R K, Sainsbury F, Lomonossoff G P, Rejzek M, Fairhurst S and Osbourn A 2013 Biochemical analysis of a multifunctional cytochrome P450 (CYP51) enzyme required for synthesis of antimicrobial triterpenes in plants *Proc. of the National Academy of Sciences* **110(35)** E3360-67

[27] Gillam M E and Hayes M 2013 The evolution of cytochrome P450 enzymes as biocatalysts in drug discovery and development *Current topics in medicinal chemistry* **13(18)** 2254-80
[28] Godiard L, Sauviac L, Dalbin N, Liaubet L, Callard D, Czernic P and Marco Y 1998 CYP76C2, an Arabidopsis thaliana cytochrome P450 gene expressed during hypersensitive and developmental cell death *FEBS letters* **438(3)** 245-49
[29] Goodwin S B and Sutter T R 2009 Microarray analysis of Arabidopsis genome response to aluminum stress *Biologia Plantarum* **53(1)** 85-99
[30] Gribskov M, McLachlan A D and Eisenberg D 1987 Profile analysis: detection of distantly related proteins *Procs. of the National Academy of Sciences* **84(13)** 4355-58
[31] Guengerich F P 2019 Cytochrome P450 research and the journal of biological chemistry *Journal of Biological Chemistry* **294(5)** 1671-80
[32] Guo P, Bai G, Carver B, Li R, Bernardo A and Baum M 2007 Transcriptional analysis between two wheat near-isogenic lines contrasting in aluminum tolerance under aluminum stress *Molecular Genetics and Genomics* **277(1)** 1-12
[33] Hamberger B, Ohnishi T, Hamberger B, Séguin A and Bohlmann J 2011 Evolution of diterpene metabolism: Sitka spruce CYP720B4 catalyzes multiple oxidations in resin acid biosynthesis of conifer defense against insects *Plant physiol* **157(4)** 1677-95
[34] Höfer R, Briesen I, Beck M, Pinot F, Schreiber L and Franke R 2008 The Arabidopsis cytochrome P450 CYP86A1 encodes a fatty acid ω-hydroxylase involved in suberin monomer biosynthesis *Journal of experimental botany* **59(9)** 2347-60
[35] Hu Y, Li W C, Xu Y Q, Li G J, Liao Y and Fu F L 2009 Differential expression of candidate genes for lignin biosynthesis under drought stress in maize leaves *Journal of applied genetics* **50(3)** 213-23
[36] Ikezawa N, Tanaka M, Nagayoshi M, Shinkyo R, Sakaki T, Inouye K and Sato F 2003 Molecular cloning and characterization of CYP719, a methylenedioxy bridge-forming enzyme that belongs to a novel P450 family, from cultured Coptis japonica cells *Journal of Biological Chemistry* **278(40)** 38557-65
[37] Irmisch S, Zeltner P, Handrick V, Gershenzon J and Köllner T G 2015 The maize cytochrome P450 CYP79A61 produces phenylacetaldoxime and indole-3-acetaldoxime in heterologous systems and might contribute to plant defense and auxin formation *BMC plant boil* **15(1)** 1-14
[38] Javid M G, Sorooshzadeh A, Moradi F, Modarres Sanavy S A M & Allahdadi I 2011 The role of phytohormones in alleviating salt stress in crop plants *Australian Journal of Crop Science* **5(6)** 726-34
[39] Ji J, Cao W, Yang L, Fang Z, Zhang Y, Zhuang M and Han F 2021 Genome-wide analysis of cabbage cytochrome P450 genes and characterization of BoCYP704B1, a gene responsible for cabbage anther development *Scientia Horticulturae* **283** 110096
[40] Jing T X, Wang D F, Ma Y P, Zeng L L, Meng L W, Zhang Q and Wang J J 2020 Genome-wide and expression-profiling analyses of the cytochrome P450 genes in Bactrocera dorsalis (Hendel) and screening of candidate P450 genes associated with malathion resistance *Pest management science* **76(9)** 2932-43
[41] Johnson S M, Lim F L, Finkler A, Fromm H, Slabas A R and Knight M R 2014 Transcriptomic analysis of Sorghum bicolor responding to combined heat and drought stress *BMC genomics* **15(1)** 1-19
[42] Jun X U, Wang X Y, Guo W Z 2015 The cytochrome P450 superfamily: Key players in plant development and defense *Journal of Integrative Agriculture* **14(9)** 1673-86
[43] Kim J E, Cheng K M, Craft N E, Hamberger B and Douglas C J 2010 Over-expression of Arabidopsis thaliana carotenoid hydroxylases individually and in combination with a β-carotene ketolase provides insight into in vivo functions *Phytochemistry* **71(2-3)** 168-78
[44] Kim Y C, Kim S Y, Paek K H, Choi D and Park J M 2006 Suppression of CaCYP1, a novel cytochrome P450 gene, compromises the basal pathogen defense response of pepper plants *Biochemical and biophysical research communications* **345(2)** 638-45
[45] Klingenberg M 1958 Pigments of rat liver microsomes *Archives of biochemistry and biophysics* **75(2)** 376-86
[46] Kong L, Anderson J M and Ohm H W 2005 Induction of wheat defense and stress-related genes in response to Fusarium graminearum *Genome* **48(1)** 29-40

[47] Kushiro T, Okamoto M, Nakabayashi K, Yamagishi K, Kitamura S, Asami T and Nambara E 2004 The Arabidopsis cytochrome P450 CYP707A encodes ABA 8′-hydroxylases: key enzymes in ABA catabolism *The EMBO journal* **23(7)** 1647-56
[48] Li H, Pinot F, Sauveplane V, Werck-Reichhart D, Diehl P, Schreiber L and Zhang D 2010 Cytochrome P450 family member CYP704B2 catalyzes the ω-hydroxylation of fatty acids and is required for anther cutin biosynthesis and pollen exine formation in rice *The Plant Cell* **22(1)** 173-90
[49] Li L, Cheng H, Gai J and Yu D 2007 Genome-wide identification and characterization of putative cytochrome P450 genes in the model legume Medicago truncatula *Planta* **226(1)** 109-23
[50] Li Y and Wei K 2020 Comparative functional genomics analysis of cytochrome P450 gene superfamily in wheat and maize *BMC plant biol* **20(1)** 1-22
[51] Liu C J, Huhman D, Sumner L W and Dixon R A 2003 Regiospecific hydroxylation of isoflavones by cytochrome p450 81E enzymes from Medicago truncatula *The Plant Journal* **36(4)** 471-84
[52] Liu Z, Boachon B, Lugan R, Tavares R, Erhardt M, Mutterer J and Renault H 2015 A conserved cytochrome P450 evolved in seed plants regulates flower maturation *Molecular plant* **8(12)** 1751-65
[53] Luo P, Wang Y H, Wang G D, Essenberg M and Chen X Y 2001 Molecular cloning and functional identification of (+)-δ-cadinene-8-hydroxylase, a cytochrome P450 mono-oxygenase (CYP706B1) of cotton sesquiterpene biosynthesis *The Plant Journal* **28(1)** 95-104
[54] Lynch T and Price A L 2007 The effect of cytochrome P450 metabolism on drug response, interactions, and adverse effects *American family physician* **76(3)** 391-96
[55] Ma B, Luo Y, Jia L, Qi X, Zeng Q, Xiang Z and He N 2014 Genome-wide identification and expression analyses of cytochrome P450 genes in mulberry (Morus notabilis) *Journal of integrative plant biology* **56(9)** 887-901
[56] Mafu S, Ding Y, Murphy K M, Yaacoobi O, Addison J B, Wang Q and Zerbe P 2018 Discovery, biosynthesis and stress-related accumulation of dolabradiene-derived defenses in maize *Plant physiology* **176(4)** 2677-90
[57] Magwanga R O, Lu P, Kirungu J N, Dong Q, Cai X, Zhou Z and Fang L 2019 Knockdown of cytochrome P450 genes Gh_D07G1197 and Gh_A13G2057 on chromosomes D07 and A13 reveals their putative role in enhancing drought and salt stress tolerance in Gossypium hirsutum *Genes* **10(3)** 226
[58] Manikandan P and Nagini S 2018 Cytochrome P450 structure, function and clinical significance: a review *Current drug targets* **19(1)** 38-54
[59] Millar A.A, Jacobsen J V, Ross J J, Helliwell C A, Poole A T, Scofield G and Gubler F 2006 Seed dormancy and ABA metabolism in Arabidopsis and barley: the role of ABA 8′-hydroxylase *The Plant Journal* **45(6)** 942-54
[60] Mizutani M 2012 Impacts of diversification of cytochrome P450 on plant metabolism Biological and *Pharmaceutical Bulletin* **35(6)** 824-32
[61] Morant M, Bak S, Møller B L, Werck-Reichhart D 2003 Plant cytochromes P450: tools for pharmacology, plant protection and phytoremediation *Current opinion in biotechnology* **14(2)** 151-62
[62] Morant M, Jørgensen K, Schaller H, Pinot F, Møller B L, Werck-Reichhart D and Bak S 2007 CYP703 is an ancient cytochrome P450 in land plants catalyzing in-chain hydroxylation of lauric acid to provide building blocks for sporopollenin synthesis in pollen *The Plant Cell* **19(5)** 1473-87
[63] Moses T, Thevelein J M, Goossens A and Pollier J 2014 Comparative analysis of CYP93E proteins for improved microbial synthesis of plant triterpenoids Phytochemistry 108 47-56
[64] Nebert D W, Nelson D R and Feyereisen R 1989 Evolution of the cytochrome P450 genes *Xenobiotica* **19(10)** 1149-60
[65] Nelson D R and Schuler M A 2013 Cytochrome P450 genes from the sacred lotus genome *Tropical plant biology* **6(2)** 138-51
[66] Nelson D R 1999 Cytochrome P450 and the individuality of species *Archives of biochemistry and biophysics* **369(1)** 1-10
[67] Nelson D R, Ming R, Alam M and Schuler M A 2008 Comparison of cytochrome P450 genes from six plant genomes *Tropical Plant Biology* **1(3)** 216-35

[68] Nelson D R, Schuler M A, Paquette S M, Werck-Reichhart D and Bak S 2004 Comparative genomics of rice and Arabidopsis. Analysis of 727 cytochrome P450 genes and pseudogenes from a monocot and a dicot *Plant Physiology* **135(2)** 756-72
[69] Nomura T, Ishihara A, Imaishi H, Endo T, Ohkawa H and Iwamura H 2002 Molecular characterization and chromosomal localization of cytochrome P450 genes involved in the biosynthesis of cyclic hydroxamic acids in hexaploid wheat *Molecular Genetics and Genomics* ***267(2)*** *210-17*
[70] Ohkawa H, Imaishi H, Shiota N, Yamada T, Inui H and Ohkawa Y 1998 Molecular mechanisms of herbicide resistance with special emphasis on cytochrome P450 monooxygenases *Plant Biotechnology* **15(4)** 173-76
[71] Ohnishi T, Watanabe B, Sakata K and Mizutani M 2006 CYP724B2 and CYP90B3 function in the early C-22 hydroxylation steps of brassinosteroid biosynthetic pathway in tomato *Bioscience, biotechnology, and biochemistry* **70(9)** 2071-80
[72] Omura T 1999 Forty years of cytochrome P450 *Biochem. Biophys. Res. Commun* **266(3)** 690-8
[73] Omura T 2018 Future perception in P450 research *Journal of inorganic biochemistry* 186 264-66
[74] Omura T and Sato R 1964 The carbon monoxide-binding pigment of liver microsomes: II. Solubilization, purification, and properties *Journal of Biological Chemistry* **239(7)** 2379-85
[75] Overkamp S, Hein F and Barz W 2000 Cloning and characterization of eight cytochrome P450 cDNAs from chickpea (Cicer arietinum L.) cell suspension cultures *Plant Science* **155(1)** 101-08
[76] Pan Z, Baerson S R, Wang M, Bajsa-Hirschel J, Rimando A M, Wang X and Duke S O 2018 A cytochrome P450 CYP 71 enzyme expressed in Sorghum bicolor root hair cells participates in the biosynthesis of the benzoquinone allelochemical sorgoleone *New Phytologist* **218(2)** 616-29
[77] Pandian B A, Sathishraj R, Djanaguiraman M, Prasad P V and Jugulam M 2020 Role of cytochrome P450 enzymes in plant stress response *Antioxidants* **9(5)** 454
[78] Park J H, Halitschke R, Kim H B, Baldwin I T, Feldmann K A and Feyereisen R 2002 A knock-out mutation in allene oxide synthase results in male sterility and defective wound signal transduction in Arabidopsis due to a block in jasmonic acid biosynthesis *The Plant Journal* **31(1)** 1-12
[79] Persans M W, Wang J and Schuler M A 2001 Characterization of maize cytochrome P450 monooxygenases induced in response to safeners and bacterial pathogens *Plant Physiol* **125(2)** 1126-38
[80] Peterson J A and Graham S E 1998 A close family resemblance: the importance of structure in understanding cytochromes P450 *Structure* **6(9)** 1079-85
[81] Prince D C, Drurey C, Zipfel C and Hogenhout S A 2014 The leucine-rich repeat receptor-like kinase BRASSINOSTEROID INSENSITIVE1-ASSOCIATED KINASE1 and the cytochrome P450 PHYTOALEXIN DEFICIENT3 contribute to innate immunity to aphids in Arabidopsis *Plant Physiol* **164(4)** 2207-19
[82] Qhanya L B, Matowane G, Chen W, Sun Y, Letsimo E M, Parvez M and Syed K 2015 Genome-wide annotation and comparative analysis of cytochrome P450 monooxygenases in Basidiomycete biotrophic plant pathogens *PLoS One* **10(11)** e0142100
[83] Rahaman M, Mamidi S and Rahman M 2018 Genome-wide association study of heat stress-tolerance traits in spring-type Brassica napus L. under controlled conditions *The Crop Journal* **6(2)** 115-25
[84] Rai A, Singh R, Shirke P A, Tripathi R D, Trivedi P K and Chakrabarty D 2015 Expression of rice CYP450-like gene (Os08g01480) in Arabidopsis modulates regulatory network leading to heavy metal and other abiotic stress tolerance *PloS one* **10(9)** e0138574
[85] Rao M J, Xu Y, Tang X, Huang Y, Liu J, Deng X and Xu Q 2020 CsCYT75B1, a citrus cytochrome P450 gene, is involved in accumulation of antioxidant flavonoids and induces drought tolerance in transgenic Arabidopsis *Antioxidants* **9(2)** 161
[86] Rasool S and Mohamed R 2016 Plant cytochrome P450s: nomenclature and involvement in natural product biosynthesis *Protoplasma* **253(5)** 1197-209
[87] Saika H, Horita J, Taguchi-Shiobara F, Nonaka S, Nishizawa-Yokoi A, Iwakami S and Toki S 2014 A novel rice cytochrome P450 gene, CYP72A31, confers tolerance to acetolactate synthase-inhibiting herbicides in rice and Arabidopsis *Plant physiol* **166(3)** 1232-40
[88] Salehi-Lisar S Y and Bakhshayeshan-Agdam H 2016 Drought stress in plants: causes, consequences, and tolerance In Drought Stress Tolerance in Plants *Springer* **1** 1-16
[89] Sanghera G S, Wani S H, Hussain W and Singh N B 2011 Engineering cold stress tolerance in crop plants *Current genomics* **12(1)** 30

[90] Schopfer C R, Kochs G, Lottspeich F and Ebel J 1998 Molecular characterization and functional expression of dihydroxypterocarpan 6a-hydroxylase, an enzyme specific for pterocarpanoid phytoalexin biosynthesis in soybean (Glycine max L.) *FEBS letters* **432(3)** 182-86
[91] Seitz C, Eder C, Deiml B, Kellner S, Martens S and Forkmann G 2006 Cloning, functional identification and sequence analysis of flavonoid 3′-hydroxylase and flavonoid 3′, 5′-hydroxylase cDNAs reveals independent evolution of flavonoid 3′, 5′-hydroxylase in the Asteraceae family *Plant molecular biology* **61(3)** 365-81
[92] Seo M and Koshiba T 2002 Complex regulation of ABA biosynthesis in plants *Trends in plant science* **7(1)** 41-48
[93] Serra O, Soler M, Hohn C, Sauveplane V, Pinot F, Franke R and Figueras M 2009 CYP86A33-targeted gene silencing in potato tuber alters suberin composition, distorts suberin lamellae, and impairs the periderm's water barrier function *Plant Physiol* **149(2)** 1050-60
[94] Stoilov I, Jansson I, Sarfarazi M and Schenkman J B 2001 Roles of cytochrome p450 in development *Drug metabolism and drug interactions* **18(1)** 33-55
[95] Stumpe M, Kandzia R, Göbel C, Rosahl S and Feussner I 2001 A pathogen-inducible divinyl ether synthase (CYP74D) from elicitor-treated potato suspension cells1 *FEBS letters* **507(3)** 371-76
[96] Sun W, Ma Z and Liu M 2020 Cytochrome P450 family: Genome-wide identification provides insights into the rutin synthesis pathway in Tartary buckwheat and the improvement of agricultural product quality *Int. Journal of Biological Macromolecules* **164** 4032-45
[97] Tabrez S and Ahmad M 2013 Cytochrome P450 system as potential biomarkers of certain toxicants: comparison between plant and animal models *Environmental monitoring and assessment* **185(4)** 2977-87
[98] Tamiru M, Undan J R, Takagi H, Abe A, Yoshida K, Undan J Q and Terauchi R 2015 A cytochrome P450, OsDSS1, is involved in growth and drought stress responses in rice (Oryza sativa L.) *Plant molecular biology* **88(1)** 85-99
[99] Tao X, Wang M X, Dai Y, Wang Y, Fan Y F, Mao P and Ma X R 2017 Identification and expression profile of CYPome in perennial ryegrass and tall fescue in response to temperature stress *Frontiers in plant science* **8** 1519
[100] Thornton L E, Rupasinghe S G, Peng H, Schuler M A and Neff M M 2010 Arabidopsis CYP72C1 is an atypical cytochrome P450 that inactivates brassinosteroids *Plant molecular biology* **74(1-2)** 167-81
[101] Trueman CL, Loewen SA, Goodwin PH 2019 Can the inclusion of uniconazole improve the effectiveness of acibenzolar-S-methyl in managing bacterial speck (Pseudomonas syringae pv. tomato) and bacterial spot (Xanthomonas gardneri) in tomato *European Journal of Plant Pathology* **155(3)** 927-42
[102] Vasav A P and Barvkar V T 2019 Phylogenomic analysis of cytochrome P450 multigene family and their differential expression analysis in Solanum lycopersicum L. suggested tissue specific promoters *BMC genomics* **20(1)** 1-13
[103] Wang E, Wang R, DeParasis J, Loughrin J H, Gan S and Wagner G J 2001 Suppression of a P450 hydroxylase gene in plant trichome glands enhances natural-product-based aphid resistance *Nature biotechnology* **19(4)** 371-74
[104] Wang M, Yuan J, Qin L, Shi W, Xia G and Liu S 2020 Ta CYP 81D5, one member in a wheat cytochrome P450 gene cluster, confers salinity tolerance via reactive oxygen species scavenging *Plant biotechnology journal* **18(3)** 791-804
[105] Wang X, Yang P, Gao Q, Liu X, Kuang T, Shen S and He Y 2008 Proteomic analysis of the response to high-salinity stress in Physcomitrella patens *Planta* **228(1)** 167-77
[106] Wang Y, Li X, Lin Y, Wang Y, Wang K, Sun C and Zhang M 2018 Structural variation, functional differentiation, and activity correlation of the cytochrome P450 gene superfamily revealed in ginseng *The plant genome* **11(3)** 170106
[107] Wedow J M, Yendrek C R, Mello T R, Creste S, Martinez C A and Ainsworth E A 2019 Metabolite and transcript profiling of Guinea grass (Panicum maximum Jacq) response to elevated [CO 2] and temperature *Metabolomics* **15(4)** 1-13
[108] Xiao F, Mark Goodwin S, Xiao Y, Sun Z, Baker D, Tang X and Zhou J M 2004 Arabidopsis CYP86A2 represses Pseudomonas syringae type III genes and is required for cuticle development *The EMBO journal* **23(14)** 2903-13

[109] Xie M M, Gong D P, Li F X, Liu G S and Sun Y H 2013 Genome-wide analysis of cytochrome P450 monooxygenase genes in the tobacco *Yi chuan= Hereditas* **35(3)** 379-87
[110] Xu F, Peng M, Luo Q, Jiang M, Zhang X, Zong X and Li Y 2015 Isolation and detection of transcript-derived fragments (TDFs) in NaCl-stressed black locust (Robinia pseudoacacia L.) using cDNA-AFLP analysis *Acta Physiologiae Plantarum* **37(8)** 1-8
[111] Yan Q, Cui X, Lin S, Gan S, Xing H and Dou D 2016 GmCYP82A3, a soybean cytochrome P450 family gene involved in the jasmonic acid and ethylene signaling pathway, enhances plant resistance to biotic and abiotic stresses *PloS one* **11(9)** e0162253
[112] Yang J, Wang G, Ke H, Zhang Y, Ji L, Huang L and Ma Z 2019 Genome-wide identification of cyclophilin genes in Gossypium hirsutum and functional characterization of a CYP with antifungal activity against Verticillium dahlia *BMC plant biology* **19(1)** 1-17
[113] Yaxley J R, Ross J J, Sherriff L J and Reid J B 2001 Gibberellin biosynthesis mutations and root development in pea *Plant physiol* **125(2)** 627-33
[114] Zeng Z, Wu Q, He B and Zeng B 2018 Genome-Wide Analysis of the Cytochromes P450 Gene Family in Aspergillus Oryzae *In Proc. of the 2018 2nd Int. Conf. on Computational Biology and Bioinformatics* 37-41
[115] Zhang M, Cao Y, Wang Z, Wang Z Q, Shi J, Liang X and Jiang C 2018 A retrotransposon in an HKT1 family sodium transporter causes variation of leaf Na+ exclusion and salt tolerance in maize *New Phytologist* **217(3)** 1161-76
[116] Zhang Y, Liu J, Zhou Y, Gong T, Wang J and Ge Y 2013 Enhanced phytoremediation of mixed heavy metal (mercury)–organic pollutants (trichloroethylene) with transgenic alfalfa co-expressing glutathione S-transferase and human P450 2E1 *Journal of hazardous materials* **260** 1100-07
[117] Zhang Y, Yu Z, Fu X and Liang C 2002 Noc3p, a bHLH protein, plays an integral role in the initiation of DNA replication in budding yeast *Cell* **109(7)** 849-60
[118] Zhang Y, Zhang B, Yan D, Dong W, Yang W, Li Q and He Z 2011 Two Arabidopsis cytochrome P450 monooxygenases, CYP714A1 and CYP714A2, function redundantly in plant development through gibberellin deactivation *The Plant Journal* **67(2)** 342-353
[119] Zheng Y, Huang Y, Xian W, Wang J and Liao H 2012 Identification and expression analysis of the Glycine max CYP707A gene family in response to drought and salt stresses *Annals of botany* 110(3) 743-756
[120] Zheng X, Li P and Lu X 2019 Research advances in cytochrome P450-catalysed pharmaceutical terpenoid biosynthesis in plants *Journal of experimental botany* **70(18)** 4619-4630

Int J Syst Assur Eng Manag
https://doi.org/10.1007/s13198-021-01153-4

ORIGINAL ARTICLE

Insights and reflections using data analytics approach of COVID-19 situation among the Indian states and ethical considerations

Zairu Nisha[1] · Sachin Kumar[2] · Saibal K. Pal[3] · Jagvinder Singh[4] · Ompal Singh[5]

Received: 21 May 2020 / Revised: 3 May 2021 / Accepted: 23 May 2021
© The Society for Reliability Engineering, Quality and Operations Management (SREQOM), India and The Division of Operation and Maintenance, Lulea University of Technology, Sweden 2021

Abstract COVID-19, a pandemic has suddenly created an unprecedented crisis in the world. COVID-19 has severely affected the lives of all the people around the world by forcing the closure or suspension of almost all socio-economic activities. Some of the countries including developed and developing such as China, the USA, Italy, Spain, the UK, France, Iran and India, etc are comprehensively affected to a deep level by this virus in terms of an incurable disease. Several countries have no options as they were hit by COVID-19 in total surprise and had no case study to follow. But some countries have attempted to face the virus such as China, Italy, Spain, Iran, etc. In this context, India has strict preventive majors in the form of a lockdown strategy besides the loosely executed testing and tracking methods. Despite all policies and decisions implementation, and their stringent execution, it is spread, and now India is in the top three COVID-19 affected countries. The rampant increase of COVID-19 cases, their response, and its impact on mortality, the effectiveness of controlling the situation by States need sharp analysis. The proposed work is an attempt to critically assess and explore the pervasive pandemic situation spread in most of the Indian state utilizing comparative analysis using data science methods contrasting them horizontally. Furthermore, the research work suggests few in-depth insights and recommends directions, steps for the public, government and other stakeholders for better management of the COVID-19 situation. Additionally, the study also deals with some ethical issues involved in the data creation, curation, and reporting on the various platforms used in the research work.

Keywords COVID-19 · Data science · Comparative analysis · Insights · Ethical considerations

✉ Sachin Kumar
official.sachin@gmail.com; skumar@cic.du.ac.in

Zairu Nisha
zairunisha@gmail.com

Saibal K. Pal
skptech@yahoo.com

Jagvinder Singh
jagvinder.singh@gmail.com

Ompal Singh
drompalsingh1@gmail.com

[1] Ramanujan College, University of Delhi, Delhi, India

[2] Cluster Innovation Centre, University of Delhi, Delhi, India

[3] Defence Research and Development Organisation, Delhi, India

[4] School of Management and Entrepreneurship, Delhi Technological University, Delhi, India

[5] Department of Operation Research, University of Delhi, Delhi, India

1 Introduction

COVID-19 is a pandemic; a disease affecting the whole world from small to large extent country after country. The first case of novel virus infection from COVID-19 was reported in China, Wuhan province in December, 2019 (Hui et al. 2019). Since then, it has spread in Europe, the USA, and Asian and Latin American Countries. COVID-19's induced situation became extremely threatening and WHO declared COVID-19 a world health emergency on 13th March, 2020 (World Health Organization et al 2020). Countries such as China, Italy, UK, Spain and Iran were

Published online: 02 June 2021

observed to be badly affected in February and early March 2020. Although In India, the first case of the COVID-19 was reported from Kerala, a southern state in the last of December 2019, the Indian government and its administration could not prepare a plan and strategy for COVID-19 till March 2020. Consequently, the Indian government decided suddenly to impose a severe and stringent lockdown on 25th March to contain the spread of the virus and break the transmission and infection chain 2020 (Sinha and Sinha 2020). There have been two observations on COVID-19 and its unstoppable infection transmission and adopted measures across the world imposing lockdown. Firstly, there has been no cure for the virus. Although scientists and researchers are working day and night to find the cure for COVID-19 in the form of vaccines. Secondly, the only way to stop the transmission of the virus is to keep a safe distance from the other person. Hence, Countries in a desperate attempt to save lives and control the COVID-19 transmission preferred the lockdown approach. Simultaneously, the administration also pushed for extensive testing of the patient through contact tracing and of those showing the symptoms of COVID-19 such as high fever, coughing, body pain etc (Cohen and Kupferschmidt 2020). The Indian government has also strictly imposed the restriction of going outside except for essential services such as food, health, law and order, and National Security Services. Due to the nature of the economy and its impact on the life of people due to the suspension of economic activities. The International Monetary Fund (IMF) and other international policy institutions have expressed great concerns regarding the job losses and contraction of the gross domestic product (GDP) of the world economy (Fernandes 2020). Few countries may be hard hit. Furthermore, few countries may go into recession and their socio-economic systems may collapse. Although it is quite clear that the economies in Europe, the USA, China, India, Japan, and Singapore are badly hit from COVID-19 (Loayza and Pennings 2020). COVID-19 has wide implications starting from economic issues such as unemployment, job losses, salary reduction, etc to personal, ethical and psychological issues about life and the future. It is not wrong to say that post corona's big implications and consequences may force us to change and adapt to the new lifestyle across the world which is sustainable for the present generation and for the future generation. Thus, India's COVID-19 situation needs detailed and thorough investigation and assessment at the state level. India is the second-largest country in terms of population, sixth largest in terms of purchasing power parity (PPP) (Mathur 2020) due to this demographic conditioning, India reported the first covid positive case in Kerala in January, whereas the rest of the country was doing well. The Indian government also observed and witnessed the severity of the situation in other affected countries which gave the policymakers and administration at least two weeks extra time to prepare and to respond effectively. The study has incorporated all the considerations to provide a wide spectrum. Following are the questions which have been addressed and discussed in depth.

1. Discussion on initiation and growth analysis of the pandemic COVID-19 in India.
2. Analysis of growth rate, fatality and recovery rate and death state wise.
3. Analysis of test-related statistics such as testing rate, infection cases, confirmed cases state wise.
4. COVID-19 transmission cases, their trends and analysis state wise.
5. Insights and recommendation for policymakers and administrators.
6. Discussion some of the ethical issues related to the data quality and data collection methods.

With this line of thought, the present paper is organised in the following manner; the first section presents the introduction. Section 2 deals with the literature survey. Section 3 explains the research methodology. Section 4 presents results and analysis. Section 5 presents insights and recommendations and Section 7 concludes the study.

2 Literature survey

An unprecedented situation posed by COVID-19 has forced not only the general public but also the scientific community, policymakers and decision-makers to think seriously and get the insights and directions based on objectivity, rationality keeping in mind the welfare of the state and the lives of people at large. Several studies have been conducted to have scientific opinion-based decision making addressing several dimensions of the COVID-19 scenario. Emanuel et al. (2020) have focused on medical resource availability and fair resource allocation. Alhazzani et al. (2020) have addressed the decision making processes and recommended some guidelines for critically ill adults to save their life and better health care. At present the real-time situation is changing rapidly, country by country are coming into the grip (Del Rio and Malani 2020). Each country has some specific characteristics depending on its population, demographic profile, geographic location, cultural aspects and development patterns. Hence the impact of COVID-19 and its intensity varies. Rodriguez-Morales et al. (2020) analysed Latin American cases with implications after the first reporting of the case in Brazil. Similarly, Gilbert et al. (2020) conducted a preparedness and vulnerability analysis of African countries. It is quite evident that COVID-19 is a new

phenomenon and no sufficient past data is available related to a pandemic with countries which they could use for their benefit and for future course of actions. As researchers were able to get the enough data to detect the COVID-19 though many advanced and latest technological developments such as AI and Machine Learning conducted experiments and presented the results to the scientific community in order to assist the decision making. Data science, Artificial Intelligence and Machine Learning have been used in many diverse fields earlier also rapidly growing and having an impact on daily life. For instance, AI have been used in the development of smart cities, energy prediction tool (Kumar et al. 2018a; Kumar and Singh 2019), intelligent transportation, especially in accident detection system (Kumar et al. 2017) and intelligent health management systems (Kumar et al. 2018b). In the COVID-19 context, researchers from China have analysed the available cases, especially in a city as Wuhan and identified some correlation of chest CT and RT-PCR testing (Ai et al. 2019). Taiwan also employed technological solutions such as data sciences, big data for better insights into the COVID-19 cases (Jason Wang et al. 2020). Arentz et al. (2020) analysed characteristics and outcomes of ill patients in Washington state and provided useful insights. Cascella et al. (2020) put forward the assessment, analysis of the treatment process of COVID-19 patients. Several studies have also been conducted to understand the patterns and trends of COVID-19 crises for course correction and proper decision making. Paquet and Schertzer (2020) discussed the COVID-19 situation as a complex Intergovernmental issue due to its pandemic nature, globalised world and cooperative and competitive decision making. Pandemic COVID-19 has many dimensions encompassing the environmental, socio-economic, ethical and political. de Paz et al. (2020) highlighted the gender dimensions of the COVID-19 pandemic. Wang et al analyse the policy of COVID-19 pandemic lockdown imposed in Chinese cities and homestay policy and its effect on COVID-19 (Wang et al. 2020). In a similar manner, Dong et al. (2020) gave a validation study for successful policy in China through meta-analysis. Lockdown has imposed restrictions on economic activities and only essential services are allowed, Mhlanga and Ndhlovu (2020) assess the Socio-economic Implications of the COVID-19 Pandemic on Livelihoods in Zimbabwe. Similarly, several studies have quantified the macroeconomic implications of COVID-19 lockdown (Mihailov et al 2020). Taking the case of India, COVID-19 and subsequent measures taken, Rajagopalan and Choutagunta (202) provided a detailed assessment of healthcare facilities and capacity in India. Vijayaraghavan and Singhal (2020) presented a comprehensive study of Indian citizens' psychological responses to COVID-19 under pandemic stringent lockdown. Lockdown due to COVID-19 has many unintended consequences such as the reduction in emission in Indian metropolitan cities, depression, lack of economic resources, premature death etc. Sharma et al. (2020) studied the effects of lockdown and its restricted emissions during COVID-19 on air quality. Gabay et al. (2020) analysed public compliance with social distancing in India. Singh and Adhikari (2020) attributed their study to analyse the impact of age on the spread, fatality, death rate and recovery in India. Samui et al. (2020) have presented a mathematical model to understand and get an insight into the COVID-19 transmission in india. Khajanchi and Sarkar (2020) worked on the forecasting of the daily and cumulative number of infections for the COVID-19. Venkateswaran and Damani (2020) assessed the effectiveness of testing, teaching and social distancing in the tracking in India and proposed a system dynamics model. Sarkar et al. (2020) also designed a model to predict the COVID-19 in Indian circumstances. Khajanchi et al. (2020) provided the dynamics of the COVID-19 pandemic in India. Observing all work done in international share and Indian domain, State-wise strategy and their patterns and gaps have not been explored by the scientific community till now. There is a lot of gap in several dimensions of the impact and assessment of COVID-19 situations in India and other states. There is no study to highlight the insights about the State-wise situation, their problems and government responses. This study addresses those gaps and puts forward the insights and recommendations based on Inter-State comparison of COVID-19 based on many trends and parameters such as morality, health infrastructure, geographic locations and their impact on their impact due to COVID-19.

3 Research methodology

The research work primarily focuses on the Indian COVID-19 scenario. India is known as a subcontinent encompassing a high level of diversity. It includes the diversity of religion, diversity of geographic and physical features, diversity of languages and dialects and one of the most important is the diversity of homogeneous and heterogeneous cultural and tradition in communities. Politically and administratively, It has been divided into political units for governance and these parts are called States and Union Territories (UT). Indian constitution provided for the division of power between union and state government by listing the subjects of governance into three lists. These lists are central list, state list and concurrent list defining the working and policymaking space for centre and state. States and central governments having their own administrative heads have taken decisions that were mainly executed by states. The COVID-19 situation including the

spread, fatality rate need thorough investigation under various parameters states wise. To answer the questions set under the light of the research objective of the study. What is the impact of their decisions, outcomes on COVID-19 situations and associated parameters such as infection rate, fatality rate, recovery rate etc needs deep analysis? This study takes a stream of data being updated on a central database state-wise and UTs wise on the parameters such as fatality rate, recovery rate, infection rate, population, health issues, infections cases, geographic locations, administrative arrangements done by states and accountable administration and development index. All these parameters taken in the form of data have been compared through data analytics tools to get useful patterns and insights. The basis of the study is to compare the actions, outcomes and situations of COVID1-9 on the basis of states and their comparative analysis, keeping states on the horizontal axis. Hence in the study, we call it horizontal analysis of the COVID-19 situation in India. More specifically, it provides insights and reports the gaps in the execution and outcomes in a particular state in comparison to other states. There are several data sources that have been merged to prepare a central database to analyse the COVID-19 scenarios in India. Data has been taken from the World health organisation(WHO) web source, John Hopkins University web source, European Union website, Indian official government web source (www.mygov.in) and www.covid-19india.org (Mygov WHO 2020). Collected data was cleaned for few missing values and collated together to get the state wise scenario. Using the data analytics approach, bar charts and multiple bar charts have been presented as visualisation tools. The presented analysis is both quantitative and qualitative. Quantitative analysis and discussion include quantitative results obtained from state-wise comparison on COVID-19 cases and population and health infrastructure criteria and doing time series analysis of the growth of COVID-19 cases in the Indian context. Qualitative analysis is presented in the discussion with the state of the health situation, immigration situation, development and growth subjective understanding state wise. The present study does not intend to model and predict the cases on COVID-19 using mathematical modelling. More importantly, the study follows a time-series pattern of events to analyse them and the spread of the COVID-19 in Indian states and their responses, approaches used by different states and their effectiveness and efficiency in containing the virus spread and infections.

4 Result and discussion

Research methodology and data visualisation tools provided directions to get insights and deep horizontal analysis of the COVID-19 situation in Indian states wise. All states and UTs have been compared on several important and relevant parameters to assess the performance and efficiency of states (having some level of autonomous decision making) in monitoring and mitigation of COVID-19 infection transmission, mortality and recovery patterns and trends. Additionally, states weaknesses and advantages have also been incorporated in the result discussions. Moreover, some demographic related data have been made the foundation to reason and support the claims as obtained and observed through data analysis. This makes better sense and gets a deep insight into the prevailing situation to make objective, rational and reliable recommendations. Figure 1 represents the time series of the COVID-19 cases reported in India in all the states and UTs starting from 25th March to 2020 to 25th April 2020, for a one-month duration. Carefully observing the plot signifies that the death rate is very less in comparison to the transmission of the COVID-19. After 25th April, COVID-19's transmission has sped up exponentially. The recovery rate is quite good. In this plot, it is observed that although initial COVID-19 cases were reported in the state, Kerala. But the number of the cases were very small by mid of April 2020. India has not seen the exponential growth of the cases or the community transmission as expected by many researchers in the predictions till the mid-week. This has been the result of stringent lockdown which stops the movement of human beings working as vectors for spreading the COVID-19. The good recovery rate has been attributed to the lifestyle of Indians and BCG vaccination in the 1960s by researchers to ensure a better immunity of the virus in the Indian population (Shivendu et al. 2020). Figure 2: gives insights about the trends and patterns of COVID-19 transmission growth in all States and UTs. The plot displays COVID-19's fatality, growth rate and recovery rate state wise. The fatality rate is high in Telangana, Andhra Pradesh, Delhi, Gujarat, Himachal, Jharkhand, Karnataka, Madhya Pradesh, Maharashtra, Meghalaya, Punjab, Rajasthan, Tamilnadu, West Bengal etc. The growth rate is high in Andhra Pradesh, Bihar, Delhi, Gujarat, Haryana, Jharkhand, Karnataka, Madhya Pradesh, Maharashtra, Odisha, Punjab, Rajasthan, Tamil Nadu, Uttarakhand, Uttar Pradesh, West Bengal and Andaman Nicobar. The recovery rate is higher in all the states. Importantly, the recovery rate is highest in states which are situated in a remote area with hilly environments and less population. This may be due to those areas where physical life is difficult which forces the cities to do physical work

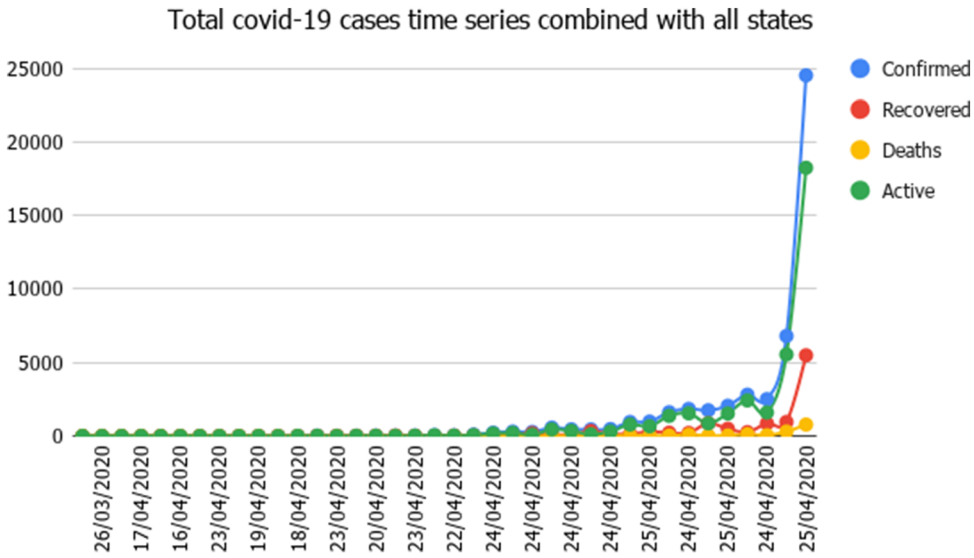


Fig. 1 The figure depicts the COVID-19 growth trends in India from March, 2020 on four type of COVID-19 cases: Confirmed COVID-19 cases, Recovered COVID-19 cases, Death due to COVID-19 and active COVID-19 cases

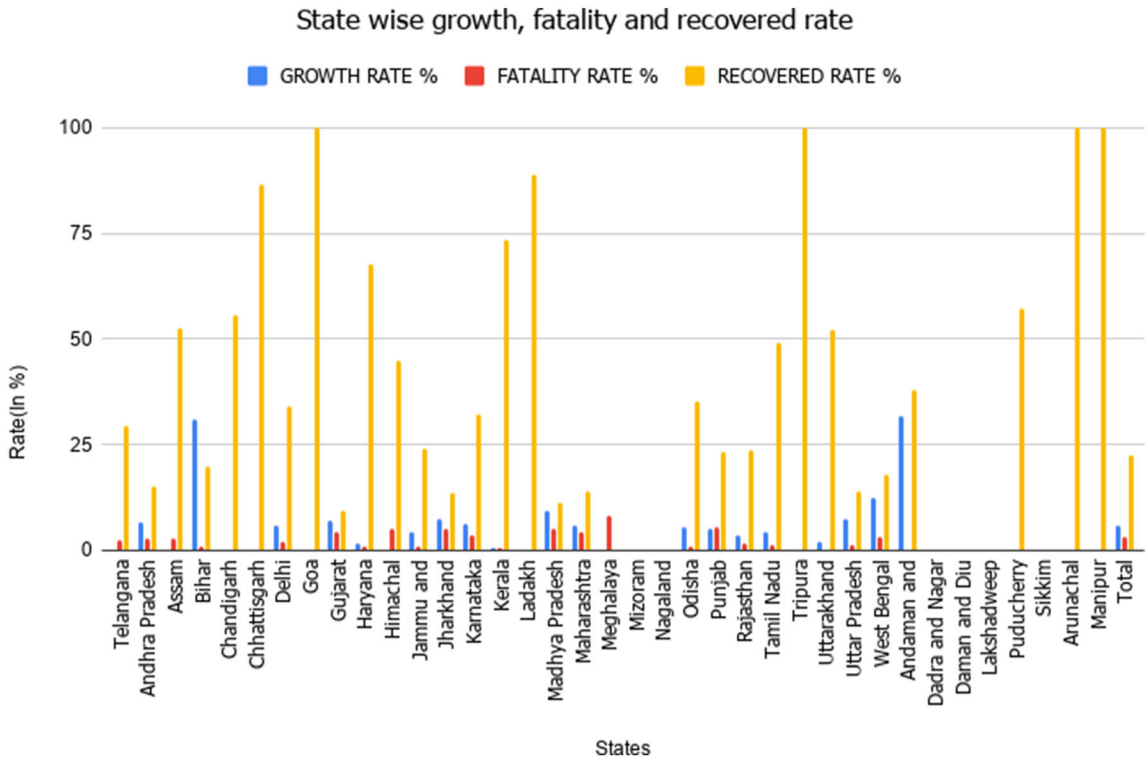


Fig. 2 Figure depicts the trends of COVID-19 growth rate in percentage, Fatality rate in percentage, Recovered rate in percentage and compare them among States and UTs Horizontally

and improve their physical strength and immunity. In contrast to this, states such as Delhi, Maharashtra, Karnataka etc with more advanced development and more population density are suffering more with high fatality and COVID-19 growth. In India, not all the states and UTs are industrialized as the development process is lopsided. Some of the states and UTs are India's engines of economic growth with development and growth in industry and service sectors. Other states such as Uttar Pradesh and Bihar have a large population seeking employment. These states labour goose to industrialised states making them migrant labourers. In this context, from the plot, it is observed that states and UTs such as Uttar Pradesh and Bihar in the beginning, contributing labourers are less affected by COVID-19. Additionally, states such as Delhi, Maharashtra, Karnataka having more advanced

transportation, industrialization and inward labour movements or migration are more COVID-19 affected and more prone to the spread of COVID-19. The industrialised States and UTs including NCT Delhi, Maharashtra, Gujarat, Karnataka. One more reason for having been affected more by COVID-19 is the international connectivity. It has been stated by many sources that testing cases are very low in India. India would have been in a better position if they would have proactively identified the traveller and cases at the airport as India has two weeks advanced observation time of what was going on in the world, especially in badly hit COVID-19 countries such as China, Italy, Iran and Spain. States having difficult terrain and geography such as Himalayan states and North-East states are better equipped due to less population, less possibility of personal contact and movement and people's social distancing behaviour. Figure 3 displays the patterns of Indian States and UTs in comparison of number of deaths, recovery rate and number of days of hospitalisation due to COVID-19. Recovery per death is highest in Kerala, followed by Haryana, Odisha, Tamil Nadu, J&K, Rajasthan and Telangana. In Delhi, it is as good although the health facilities are better as compared to other states and UTs. Why is Delhi suffering more? There may be many reasons. International flights and commuting citizens, densely populated UTs and high levels of population in most of the time of year. Pollution reduced immunity and reduced up to 5 years of life. Recovery per death tells that more people recover and fewer are dying after the infection. Haryana is a considered physically healthy State of India due to its food habits and physical strengths and farming occupation. The best recovery rate is in the Himalayan States such as Uttrakhand, Jammu and Kashmir, Himachal Pradesh and north-eastern states as compared to states having plain geography. States having plain geography and falling in between the Yamuna and the Ganga have huge populations, high levels of water and air pollution, lack of health cards and educational facilities due to a high degree of competition with fewer resources and corruption. The highest days of hospitalisation have been observed in the Ladakh region due to the extreme weather of the region. In most of the States and UTs, the hospitalisation days are similar that is from 10 days to 15 days. Figure 4 presents the growth of COVID-19 infections in terms of the rate of doubling day wise. More clearly, it is the number of days in doubling the number of COVID-19 cases. This is a very important parameter. During the linear growth, this parameter is less but as the growth rate takes the shape of the exponential function, it reduced drastically. This happens during the community transmission stage when there is no control on the contact tracing and infection transmission and the only way left is to upgrade the health infrastructure and respond to the uncontrollable situation. To prevent this situation, Lockdown has been proposed and implemented. The plot gives the impression that the doubling rate is high in Delhi, Gujarat, Maharashtra, Tamil

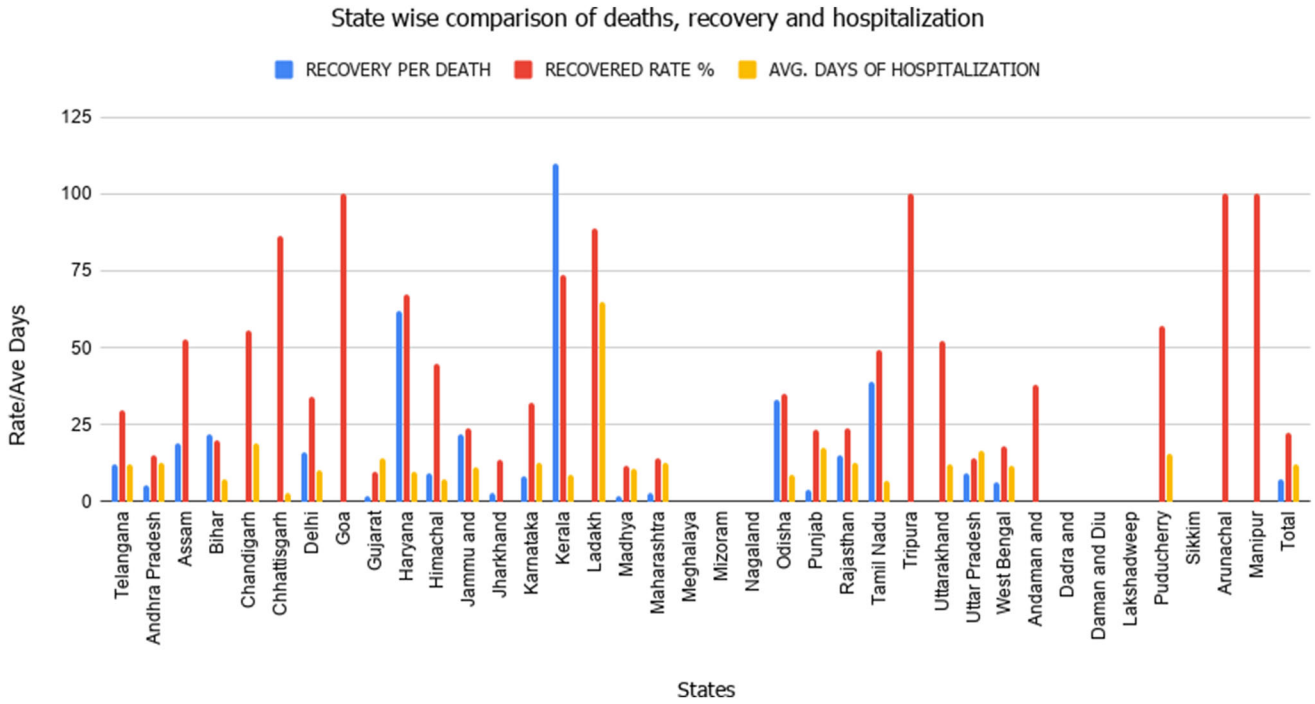


Fig. 3 Figure depicts the trends of COVID-19 recovery per death, recovered rate in percentage and number of days of hospitalisation and compare them among States and UTs Horizontally

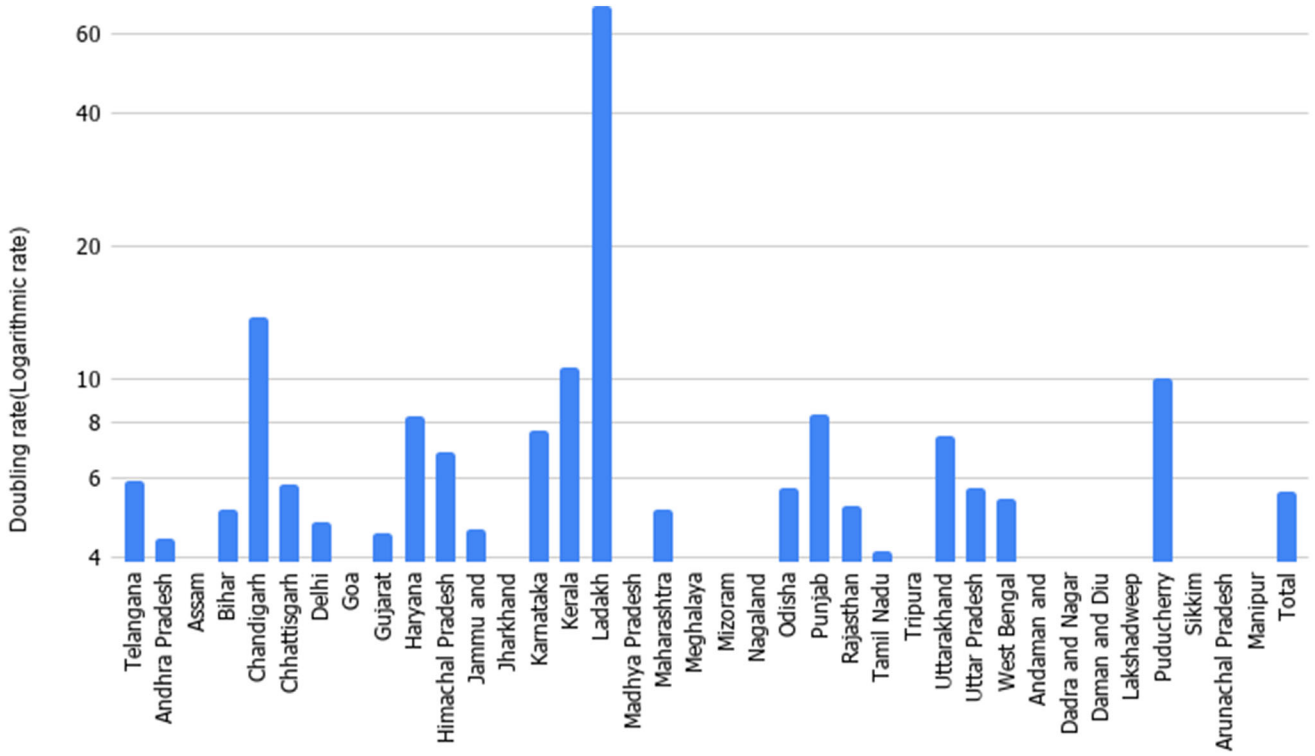


Fig. 4 Figure depicts the trends of COVID-19 doubling rate (number of days taken in doubling of active COVID-19 cases) and compare them among States and UTs Horizontally. Data is plotted in logarithmic scale

Nadu and Andhra Pradesh. All these states have substantial levels of industrialisation and migration from other states. These states also have more urbanisation as compared to the other states which have shown a low doubling rate due to their difficult terrain and less economic development. Such observations have inferred that more industrialised and urbanised regions must have better health and residential services to avoid the disease. States which are geographically difficult are better at having a high doubling rate. Here regional planning may also come into the picture if the case of Chandigarh is taken which has quite a high doubling rate and the city is a well developed and urban region of the State, Punjab.

Figure 5 puts forward the trends of the testing strategy and methods adopted by each State and UT. The figure compares states and UTs on tests per million, positive cases per million and per cent of confirmed over total tests. It is observed and reported in international media that India's testing rate is very less as compared to other countries taking into consideration its population (IANS 2020). It is even less than that of Pakistan (Mygov WHO 2020). The testing rate measures the number of COVID-19 test cases conducted per million population. In India, Delhi is leading in terms of number of tests per million cases. Delhi is considered a good state in terms of health facilities due to the priority of the Delhi government which has allocated approximately 20 per cent of the economic budget to health and introduced several acclaimed new health facilities and schemes such as new hospitals and decentralisation of health facilities through PolyClinic and Mohalla Clinic. This is one of the reasons for having more cases as more testing is being done. This is reflected in reporting of positive cases as that number is also higher in Delhi. The other States and UTs are facing a problem of lack of testing kits and lack of trained health professionals (Shekhar 2020). They are more focused on stringent lockdown imposition strategies and social distancing. Although, it seems that Lockdown is not the final answer. As soon as the lockdown is lifted and people having COVID-19 cases which are also asymptomatics start travelling, COVID-19 infections will spread in the second wave. Experts report that second waves of COVID-19 may be more fatal than this phase. So more testing must be done and more resources should be allocated to contain the spread and stop the community transmission stage. Figure 6 compares the results of the several states and UTs on four very important points related to COVID-19. Four parameters are the percentage share of states and UTs in total COVID-19 infections/cases, the share of states and UTs in total COVID-19 active cases, the share of states and UTs in total COVID-19 death cases, the share of states and UTs in COVID-19 tests. From the plot it is clear that the

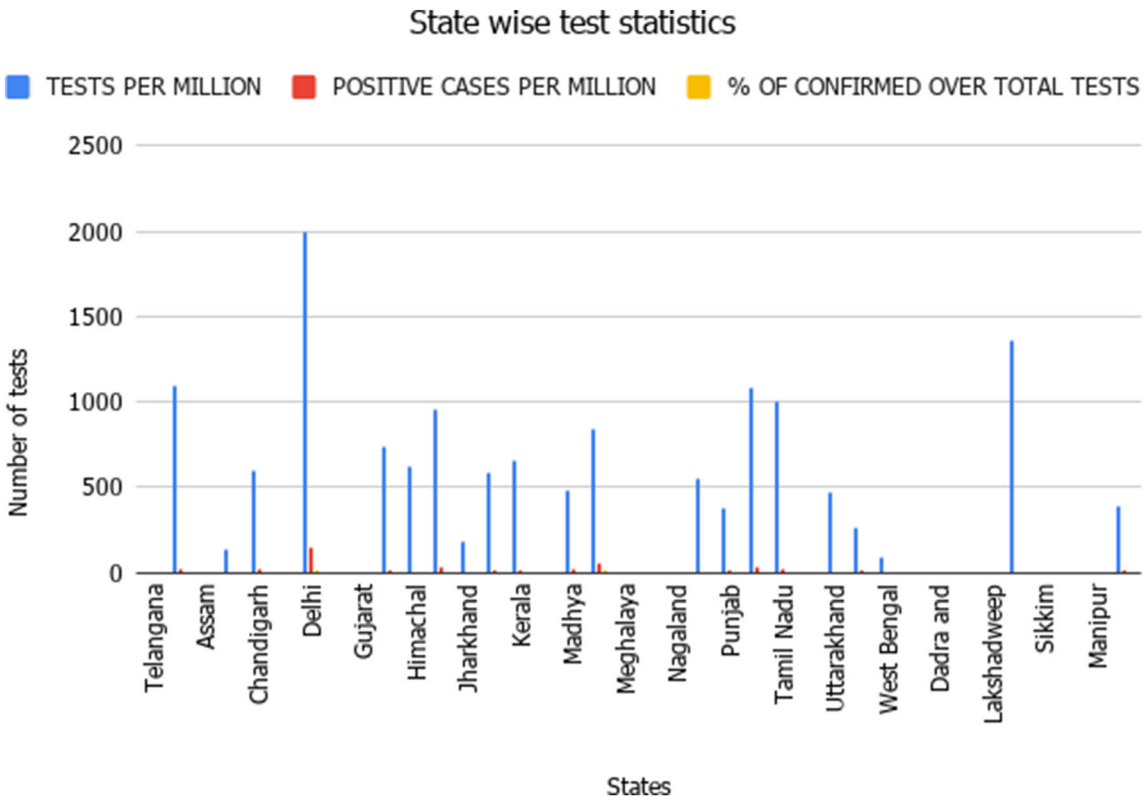


Fig. 5 Figure depicts the trends of COVID-19 test patterns and statistics and compare them among States and UTs horizontally

share of the rate of mortality and share COVID-19 death cases is highest in Maharashtra, followed by Gujarat and Delhi and Madhya Pradesh(MP). Interestingly, Each of the states and UTs has different reasons for the spread of COVID-19 and death due to COVID-19. Maharashtra and Gujarat follow similar trends and reasons for COVID-19 as both of them have an almost similar level of industrialisation and urbanisation. Decision making has been in confusion due to change in the administration and political authority. This situation forced the delay in active decision making in MP. Specifically, the top level of bureaucracy also suffered from COVID-19 and their health workers of top-level also faced similar problems forcing them and the system to complete breakdown. NCT Delhi has been able to manage its death rate to thirty due to the proactive testing rate and proactive measures taken by the government and volunteers. The figure depicts that the deaths are mostly in these four states and UTs. Another observed situation from our data visualisation is that in states like Gujarat and Maharashtra, their rate of testing is not near the satisfactory level. Good testing rate of COVID-19 helps in better health care provisions, mitigation and recovery than the lack of test cases which only contributes to more virus spread but also increases the mortality rate. In the case of the share of the COVID-19 cases, Maharashtra and Gujarat are leading followed by Delhi and Maharashtra. This is depicted in Fig. 7, also with state and UTs wise map of India. The reasons again are the same as States have more congested living and lack of health and other facilities for migrant workers from other poor states, part of the country. There is an interesting note here that there is no or very few active cases of COVID-19 in Chhattisgarh, Goa, Mizoram, Nagaland Meghalaya, Ladakh, Tripura, Uttarakhand, Pondicherry and Sikkim etc. All these states either have better life expectancy and immunity or better way of living and also better education and awareness about the issues. These states are also from mountainous regions.

5 Insights and recommendations

After discussion and critical analysis of the results, this section suggests some measures which are better for decision making and administration for government officials and policymaking. States which are doing good are NCT Delhi and other remote areas such as the Himalayan States and Northeastern states including Kerala and Goa. Other states need to learn from their strategies, especially political and administrative decision making. Lockdown is a good measure but not a sufficient measure. Lockdown needs the support of extensive testing support to ensure the tracing and quarantine of affected persons. Imposing

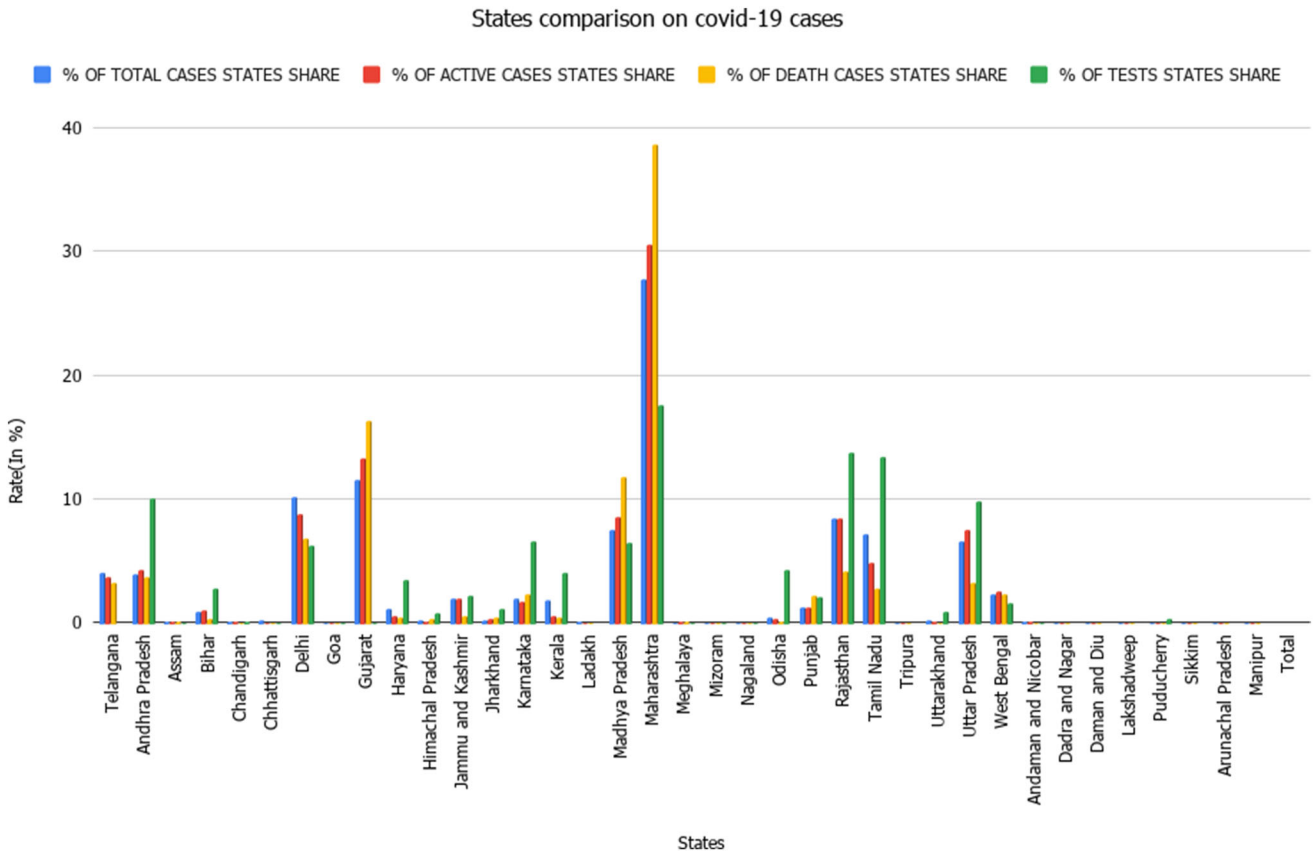


Fig. 6 Figure depicts the trends of COVID-19 percentage of state in total cases, percentage of state in active cases, percentage share of death cases and percentage share of test cases and compare them among States and UTs horizontally

lockdown for a longer period will surely have a bad impact on the large population of this country, especially on poor people, daily wages earners, farmers and small shopkeepers as they don't have enough food and money to support their livelihood. District wise strategy should be adopted and locking of districts with clear and unambiguous guidelines keeping future circumstances and situations in mind, needs to be followed. Better educated states and UTs are better in health facilities and consequently better equipped and proactive in dealing with the pandemic. Hence investment in human capital is very much important and needed. Better education and health promote cleanliness, awareness and hygiene which improves the immunity of human beings and ultimately result in better and fewer chances of disease. It has been observed that the death rate is higher in the older population (Dowd et al. 2020) which is greater than 60 years so a better and targeted approach should be used about them and allow the younger population for the economic activity to run the cyclone of development.

6 Ethical issues and considerations

COVID-19 transmission and infection in the community have triggered collaborative efforts at the individual level, community level, institutional level, government level and intergovernmental levels. Technological giants and individuals are contributing to the data and its collection methods and in the development of analytical tools in order to understand the deep pattern of transmission and spread, demographic attribute's relationships with COVID-19 infections, relationship with medical health history, current location, movements of infected people. All this is being done through data collections systems such as websites and apps. Contact tracing apps such as the Arogya Setu app by the Indian Government, have been pushed to be installed for keeping safe the healthy person from the infected person by informing them of the rule-based quantified risk exposure in the surroundings of individuals. There are many individual efforts whose data validity and reliability are questioned. Contact tracing Apps are one of digital surveillance in the pandemic which needs proper protection of individual data and medical history and attributes from going in the hand of adversaries. In addition to this, in the

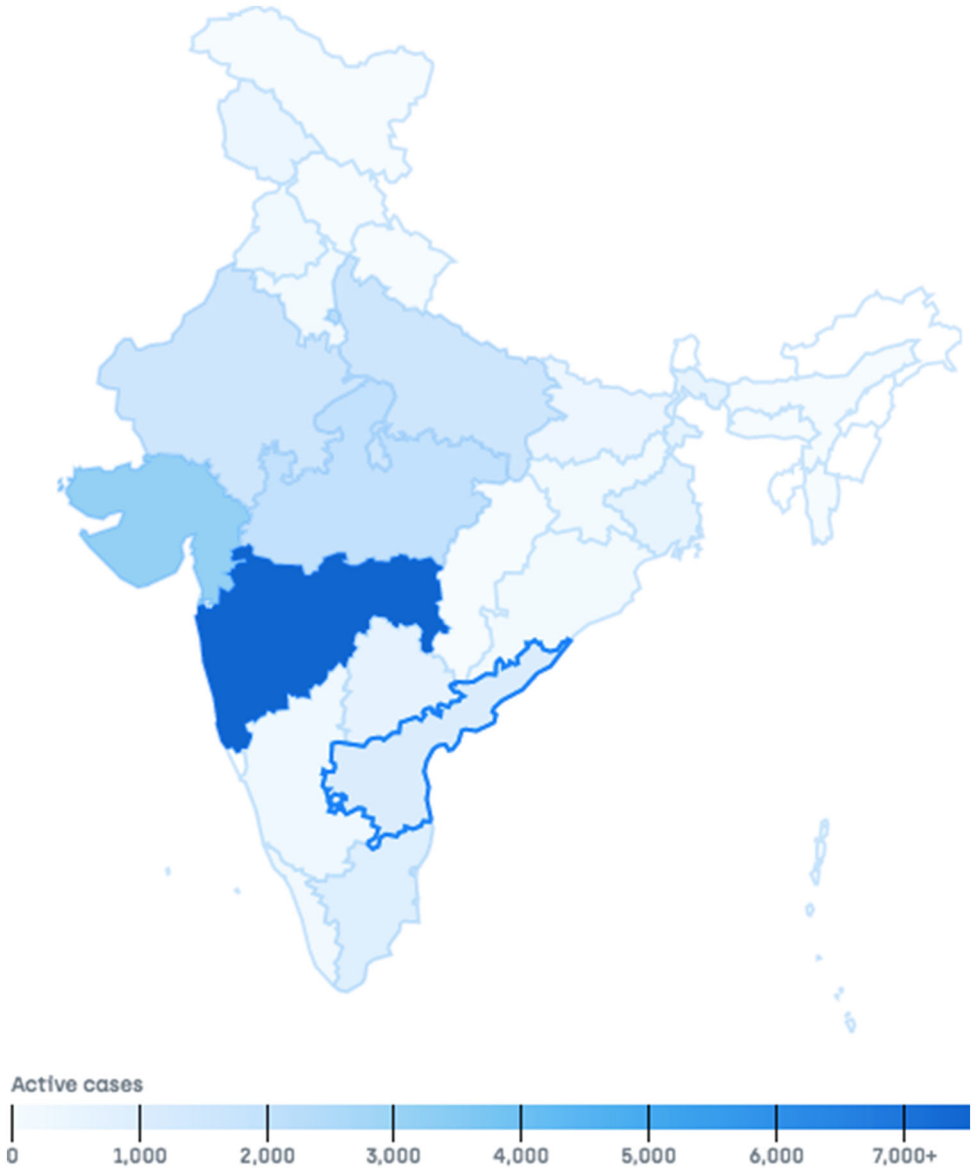


Fig. 7 State wise map of COVID-19 infection cases. Dark color represents more COVID-19 cases and light color represents less number of cases (Mygov WHO 2020)

present circumstances as posed by COVID-19, the Digital technologies and analytics framework of prediction and data collection should include ethical data collection practices, consent taking steps in all their data collection methods and stages. Digital technology tools accessing data need to conform to the international guidelines of data collection and access and storage of information without compromising the security and privacy of individuals. Ethical practices related to data sharing should also be considered so that the community trusts the finding of the research work. There is a need for clarity of perspective on the rapidly evolving field of human data collection using digital tools. The ethical and legal boundaries of such tools for disease surveillance and contact tracing are unclear. This needs clarity, equitable and egalitarian policies for the stakeholders.

7 Conclusion

This proposed research work has done extensive analysis of the information related to Indian states and UTs and their adopted measures and strategies and outcome of those strategies. Results and discussion shows that States with higher education and better health services and proactive leadership are performing better in mitigation and containment of COVID-19 spread. States with more industrialised economies and labour in or out from the other states are lagging behind and suffering more from COVID-19. States with better law and order with stringent rules and regulation and their execution are also doing better. Geologically different terrain states are also in a better position to handle the COVID-19 situation. There is a need to learn from better-performing states and more focus should be

given to testing rate and improving health services in states. Additionally, data collection and disease surveillance tools should be ethically and legally conforming to the guidelines and framework of the international community and stakeholders.

References

Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, Tao Q, Sun Z, Xia L (2019) Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. Radiology 296:200642

Alhazzani W, Møller MH, Arabi YM, Loeb M, Gong MN, Fan E, Oczkowski S, Levy MM, Derde L, Dzierba A (2019) Surviving sepsis campaign: guidelines on the management of critically ill adults with coronavirus disease 2019 (COVID-19). Intensive Care Med 46(1–34):2020

Arentz M, Yim E, Klaff L, Lokhandwala S, Riedo FX, Chong M, Lee M (2020) Characteristics and outcomes of 21 critically ill patients with COVID-19 in Washington state. Jama 323:1612–1614

Cascella M, Rajnik M, Cuomo A, Dulebohn SC, Di Napoli R (2020) Features, evaluation and treatment coronavirus (COVID-19). In: Statpearls [internet]. StatPearls Publishing

Cohen J, Kupferschmidt K (2020) Countries test tactics in 'war' against covid-19

de Paz C, Muller M, Maria Munoz Boudet A, Gaddis I (2020). Gender dimensions of the COVID-19 pandemic

Del Rio C, Malani PN (2020) Covid-19—new insights on a rapidly changing epidemic. Jama 323:1339–1340

Dong J, Zhou Y, Zhang Y, Franz D (2020) A validation study for the successful isolation policy in china: a meta-analysis in COVID-19. medRxiv

Dowd JB, Andriano L, Brazel DM, Rotondi V, Block P, Ding X, Liu Y, Mills MC (2020) Demographic science aids in understanding the spread and fatality rates of COVID-19. Proc Natl Acad Sci 117(18):9696–9698

Emanuel EJ, Persad G, Upshur R, Thome B, Parker M, Glickman A, Zhang C, Boyle C, Smith M, Phillips JP (2020) Fair allocation of scarce medical resources in the time of COVID-19

Fernandes N (2020) Economic effects of coronavirus outbreak (COVID-19) on the world economy. Available at SSRN 3557504

Gabay G, Bellissimo N, Garby I, Zemel G, Gere A, Moskowitz HR, et al (2020) Driving public compliance with the health policy of social distancing: the case of the COVID-19 pandemic in India

Gilbert M, Pullano G, Pinotti F, Valdano E, Poletto C, Boëlle P-Y, d'Ortenzio E, Yazdanpanah Y, Paul Eholie S, Altmann M et al (2020) Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study. The Lancet 395(10227):871–877

Hui David S, Azhar Esam I, Madani Tariq A, Ntoumi F, Kock R, Dar O, Ippolito G, Mchugh Timothy D, Memish Ziad A, Drosten C et al (2019) The continuing 2019-ncov epidemic threat of novel coronaviruses to global health-the latest novel coronavirus outbreak in Wuhan, China. Int J Infect Dis 91(264):2020

IANS (2020) India's covid testing rate is low, says who chief scientist. The Times of India

Khajanchi S, Sarkar K (2020) Forecasting the daily and cumulative number of cases for the COVID-19 pandemic in India. Chaos Interdiscip J Nonlinear Sci 30(7):071101

Khajanchi S, Sarkar K, Mondal J, Perc M (2020) Dynamics of the COVID-19 pandemic in India. arXiv:2005.06286

Kumar S, Kalia A, Sharma A (2017) Predictive analysis of alertness related features for driver drowsiness detection. In: International conference on intelligent systems design and applications. Springer, pp 368–377

Kumar S, Saibal Pal K, Singh R (2018a) A novel method based on extreme learning machine to predict heating and cooling load through design and structural attributes. Energy Build 176:275–286

Kumar S, Pal SK, Singh RP (2018b) A conceptual architectural design for intelligent health information system: case study on India. In: Quality, IT and Business Operations. Springer, pp 1–15

Kumar S, Pal SK, Singh R (2019) A novel hybrid model based on particle swarm optimisation and extreme learning machine for short-term temperature prediction using ambient sensors. Sustain Cities Soc

Loayza NV, Pennings S (2020) Macroeconomic policy in the time of COVID-19: a primer for developing countries

Mathur AN (2020), Strategies for solving wicked problems of true uncertainty: tackling pandemics like covid-19 (version: April 13, et al (2020) Technical report. Research and Publication Department, Indian Institute of Management Ahmedabad, p 2020

Mhlanga D, Ndhlovu E (2020) Socio-economic implications of the COVID-19 pandemic on smallholder livelihoods in Zimbabwe

Mihailov A et al (2020) Quantifying the macroeconomic effects of the COVID-19 lockdown: Comparative simulations of the estimated galí-smets-wouters model. Technical report, Henley Business School, Reading University

Mygov WHO, John Hopkins University (2020) Data sources for COVID-19

Paquet M, Schertzer R (2020) Covid-19 as a complex intergovernmental problem. Can J Polit Sci 53:1–5

Rajagopalan S, Choutagunta A (2020) Assessing healthcare capacity in India

Rodriguez-Morales AJ, Gallego V, Escalera-Antezana JP, Mendez CA, Zambrano LI, Franco-Paredes C, Suárez JA, Rodriguez-Enciso HD, Balbin-Ramon GJ, Savio-Larriera E et al (2020) Covid-19 in Latin America: the implications of the first confirmed case in Brazil. Travel Med Infect Dis

Sarkar K, Khajanchi S, Nieto JJ (2020) A mathematical model for COVID-19 transmission dynamics with a case study of India. Chaos Solitons Fract 140:110173

Kankan S, Subhas K, Nieto Juan J (2020) Modeling and forecasting the COVID-19 pandemic in India. Chaos Solitons Fract 139:110049

Sharma S, Zhang M, Gao J, Zhang H, Kota SH et al (2020) Effect of restricted emissions during COVID-19 on air quality in India. Sci Total Environ 728:138878

Shekhar S. A quantitative framework for expanding testing capacity of COVID-19 in India

Shivendu S, Chakraborty S, Onuchowska A, Patidar A, Srivastava A (2020) Is there evidence that BCG vaccination has non-specific protective effects for COVID-19 infections or is it an illusion created by lack of testing? Available at SSRN 3579847

Singh R, Adhikari R (2020) Age-structured impact of social distancing on the COVID-19 epidemic in India. arXiv:2003.12055

Sinha S, Sinha N (2020) Modelling of covid—19 infection propagation data in Indian population

Venkateswaran J, Damani O (2020) Effectiveness of testing, tracing, social distancing and hygiene in tackling COVID-19 in India: a system dynamics model. arXiv:2004.08859

Vijayaraghavan P, Singhal D (2020) A descriptive study of indian general public's psychological responses during COVID-19 pandemic lockdown period in India

Wang CJ, Ng CY, Brook RH (2020) Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing. Jama 323:1341–1342

Wang J, Tang K, Feng K, Lv W. When is the COVID-19 pandemic over? Evidence from the stay-at-home policy execution in 106 Chinese cities

World Health Organization et al (2020) Who director-general's opening remarks at the media briefing on COVID-19—11 March 2020. Geneva, Switzerland

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

symmetry

MDPI

Article

Intra-Block Correlation Based Reversible Data Hiding in Encrypted Images Using Parametric Binary Tree Labeling

Arun Kumar Rai [1], Neeraj Kumar [2], Rajeev Kumar [3], Hari Om [1], Satish Chand [4] and Ki-Hyun Jung [5,*]

[1] Department of Computer Science and Engineering, Indian Institute of Technology (Indian School of Mines), Dhanbad 826004, India; arun.18DP000380@cse.iitism.ac.in (A.K.R.); hari@iitism.ac.in (H.O.)
[2] Department of Electronics and Communication Engineering, Nagarjuna College of Engineering and Technology, Bengaluru 562164, India; neeraj.mohiwal@gmail.com
[3] Department of Computer Science and Engineering, Delhi Technological University, Delhi 110042, India; rajeevkumar@dtu.ac.in
[4] School of Computer and System Sciences, Jawaharlal Nehru University, Delhi 110067, India; schand@mail.jnu.ac.in
[5] Department of Cyber Security, Kyungil University, Kyeongbuk 38424, Korea
* Correspondence: kingjung@kiu.kr or khanny.jung@gmail.com; Tel.: +82-53-600-5626

Abstract: In this paper, a high capacity reversible data hiding technique using a parametric binary tree labeling scheme is proposed. The proposed parametric binary tree labeling scheme is used to label a plaintext image's pixels as two different categories, regular pixels and irregular pixels, through a symmetric or asymmetric process. Regular pixels are only utilized for secret payload embedding whereas irregular pixels are not utilized. The proposed technique efficiently exploits intra-block correlation, based on the prediction mean of the block by symmetry or asymmetry. Further, the proposed method utilizes blocks that are selected for their pixel correlation rather than exploiting all the blocks for secret payload embedding. In addition, the proposed scheme enhances the encryption performance by employing standard encryption techniques, unlike other block based reversible data hiding in encrypted images. Experimental results show that the proposed technique maximizes the embedding rate in comparison to state-of-the-art reversible data hiding in encrypted images, while preserving privacy of the original contents.

Keywords: image encryption; reversible data hiding; parametric binary tree labeling; privacy; intra-block correlation

check for updates

Citation: Rai, A.K.; Kumar, N.; Kumar, R.; Om, H.; Chand, S.; Jung, K.-H. Intra-Block Correlation Based Reversible Data Hiding in Encrypted Images Using Parametric Binary Tree Labeling. *Symmetry* **2021**, *13*, 1072. https://doi.org/10.3390/sym13061072

Academic Editors: Kuo-Hui Yeh and Peng-Yeng Yin

Received: 20 April 2021
Accepted: 7 June 2021
Published: 16 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Cloud computing is a great technology for remote sharing and accessing of information across the world. It renders a simplistic model for real-time information sharing, accessing, and/or storing on the cloud server [1] on a pay-per-use basis. Due to its simplicity, today, cloud computing services have millions of subscribers who store their public/private data on cloud servers. The cloud servers possess several data storage devices for storing subscribers' data which can be any multimedia data (such as image, video, audio), text data, program data, etc., in any format. To protect against information leakage and maintain confidentiality, subscriber data is first encrypted and then stored on the cloud servers [2]. To manage the cloud services and data storages, and also to control loading–unloading of data between storages for load balancing, the cloud servers makes use of a cloud administrator which attaches some auxiliary information, such as subscriber name, file name, file type, source information, etc., to subscriber's data for its efficient handling without decoding the subscriber data. The auxiliary information is attached secretly by embedding the same in the encrypted subscriber data so that unauthorized access can be prevented without incurring additional storage cost. Since the auxiliary information is cloud management information, it is embedded in a lossless and reversible manner in the subscriber data. For this, there exist several reversible data hiding (RDH) techniques

which have been developed in the past for lossless embedding and recovery of both secret data and subscriber data. Some of the popular RDH techniques are lossless compression based [3–6], difference expansion-based [7], histogram expansion, and prediction error expansion based [8–12] RDH techniques. However, these techniques are only suitable for plaintext data, but not for encrypted data.

In 2008, Puech et al. [13] unveiled an RDH technique for encrypted images (RDHEI). They first encrypted the original image using Advanced Encryption Standard (AES) and then embedded the secret data bits at random locations in each 4×4 by simple substitution. At the receiver end, local standard deviation analysis is performed to reinstate the original image and get back the hidden secret data. After Puech et al.'s technique [9], the RDHEI field has been explored by various researchers [14–36].

Based on the analysis of available RDHEI literature, the RDHEI techniques can be classified into three major categories, namely, vacating room after encryption (VRAE) [12,15,21,22], vacating room by encryption (VRBE) [32,34,35], and reserving room before encryption (RRBE) [23,29–31,33]. The VRAE techniques create vacancies for secret data hiding in encrypted data whereas techniques that fall into the VRBE category create vacancies for secret data hiding during encryption. As far as the RRBE category is concerned, first, a room is reserved in the unencrypted media which is then exploited for embedding the secret data after the media's encryption. Among these three approaches, the VRAE approach offers very limited embedding capacity (EC) because encrypted data provides limited room for secret data hiding as encryption destroys the correlations of subscriber data. Therefore, the cloud administrator is unable to produce many vacancies for secret data embedding in encrypted data. In the VRBE approach, specific encryption tactics are applied to subscriber data, which maintain some correlations in encrypted data. Later, the correlations present in the encrypted data are exploited by the cloud administrator for embedding auxiliary information. Still, this approach does not yield a high EC as spatial redundancy existing in the subscriber data is not fully utilized. However, the RRBE approach is an altogether different approach, in which subscriber data in unencrypted form is exploited to reserve vacancies. This means that, later on, the subscriber data can be encrypted and auxiliary information can be embedded in the encrypted data, by exploiting the redundancy of non-encrypted data due to the reserved vacancies. Thus, the RRBE approach renders higher EC than the VRAE and VRBE approaches.

In RDHEI, the encryption of plain-text data is performed by the subscriber using an encryption key K_e, who then uploads the encrypted data on the cloud server. Next, the cloud administrator performs embedding of auxiliary information using a secret key K_s to transform the encrypted data into marked encrypted data. The marked encrypted data accommodates both subscriber data and auxiliary information in encrypted form, and is stored on the cloud server. At the decoder side, three cases may exist. In the first case, if a user owns both the encryption key K_e and the secret key K_s, he/she can retrieve both subscriber data and auxiliary information. In case only the encryption key K_e or the secret key K_s is possessed by the user then only subscriber data or the auxiliary information can be recovered, respectively. A schematic diagram of the RDHEI approach is shown in Figure 1.

In the literature, some RDHEI techniques have been introduced [14–18] which allow retrieval of auxiliary information only after decryption of marked encrypted data. In other words, auxiliary information can be extracted using a secret key K_s only after performing the decryption using the encryption key K_e. TheseRDHEI techniques are known as non-separable techniques. In [19–35], another variant of RDHEI techniques, i.e., separable RDHEI techniques, was introduced to overcome the limitation of non-separability. These techniques allow separate restoration of subscriber data and extraction of auxiliary information from the marked encrypted information. In other words, an encryption key K_e keeper can perform the decryption process separately without being restricted by the auxiliary information extracting process, and similarly a secret key K_s keeper can perform the auxiliary information extracting process without being restricted by the decryption process.

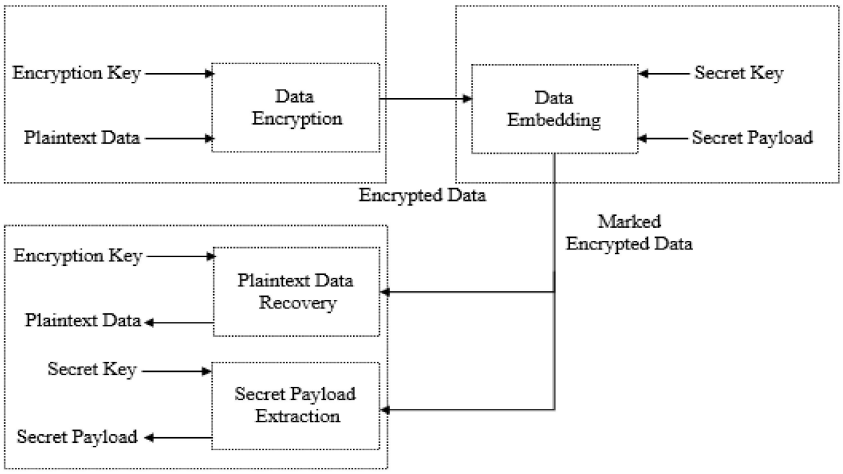


Figure 1. A schematic diagram of RDHEI technique.

In this paper, an intra-block correlation based high capacity RDHEI technique using PBTL is proposed. In the proposed technique, the original image is first partitioned into non-overlapping blocks of a predetermined size (e.g., 4×4) which are then categorized into smooth, moderately complex, and highly complex blocks. Now, prediction errors and residual errors are calculated for pixels of the smooth blocks and moderately complex blocks. Next, the proposed technique applies a stream encryption method (different from Yi et al.'s technique) to the original host image to improve the security of the encrypted subscriber data (or the encrypted host image). The pixels of the encrypted image are then labeled using the PBTL scheme and the calculated prediction and residual errors. Finally, the labelled pixels are exploited to reversibly embed the secret data. Contributions of the proposed RDHEI technique can be summarized as follows:

(1) The proposed work discloses a high capacity RDHEI technique using PBTL. The proposed RDHEI technique first partitions the original image into uniformly sized blocks and then categorizes the image blocks into three categories, i.e., smooth, moderately complex, and highly complex blocks based on the prevalent correlation of each block. This granular classification of the image blocks helps in reserving pixels for efficiently embedding the secret information.
(2) The smooth and moderately complex blocks generally possess high and decent pixel correlation, respectively. Thus, both types of blocks are used for embedding the secret information whereas the highly complex blocks are snubbed in the data embedding process as there is no or minimal correlation.
(3) Further, the proposed technique uses a stream encryption method for encrypting the original image. The stream cipher provides higher security in comparison to block cipher as it completely destroys the correlation of image pixels.
(4) In addition, the proposed technique is a separable RDHEI technique which allows separate recovery and extraction of subscriber data and the secret message in a lossless manner.
(5) Experimental results show that the proposed RDHEI technique has superior embedding performance in comparison to related previous RDHEI techniques while ensuring security of the image contents.

The rest of this paper is structured as follows. Section 2 discusses related separable RDHEI techniques. Section 3 introduces the review of parametric binary tree labeling (PBTL) scheme. The proposed RDHEI technique is discussed in Section 4. Section 5 shows the experimental results and their comparative analysis. Lastly, a conclusion as well as the scope of future works has been drawn for the proposed work.

2. Related Works

In this section, some of the popular RDHEI techniques are briefly reviewed. In 2012, Zhang discussed a separable RDHEI technique [19]. In the separable RDHEI technique [19], the subscriber data (which is an image) is first encrypted using a standard encryption method and then a room/space is reserved for embedding the secret/auxiliary information by compressing three least significant bits (LSBs) of the encrypted data. However, the generated space is sparse in nature which limits the embedding rate. Xiaotian et al. [17] discussed a separable RDHEI technique which outperforms the earlier technique [19]. Dragoi et al. [21] proposed a new separable RDHEI scheme by creating space after the encryption. The main feature of this scheme is the use of a two-stage data hiding process in which the reference pixel is predicted based on the median context value. As far as performance improvement is concerned, the proposed scheme marginally improves embedding capacity in comparison to erstwhile related schemes. In 2018, Dragoi et al. [22] again came up with a new separable RDHEI scheme with color images, where correlation between RGB color planes of a target pixels and its correlation with neighboring pixels are exploited to embed data in a color image. In 2013, Ma et al. unfolded the first RRBE approach based RDHEI technique [23]. The technique partitions the image to get smooth and complex regions. Next, one or more LSBs of complex regions are embedded into the smooth regions using any of the conventional RDH algorithms such as [5–7] to create room in the rough regions. Finally, the reserved room in the encrypted images are filled by auxiliary information. This reservation of space in the complex regions yields a high embedding rate which goes up to 0.5 bit-per-pixel (bpp). Similarly, Zhang et al. reserved room for secret data hiding using the prediction error (PE) based histogram shifting method [18]. In [24], Xu et al. disclosed the calculation of prediction error based on interpolation technique then applied histogram shifting and the difference expansion technique to exploit the prediction error for data hiding. However, only a small improvement in performance was achieved. In [25], Mathew et al. refine the work of Ma et al. [23] by introducing a new pixel intensity variation criterion for classifying image blocks into smooth blocks and rough blocks. However, the work only marginally improves the embedding rate. To further upgrade the embedding rate, Cao et al. [26] discuss a patch-level sparse representation method for RDHEI. The method performs encryption in phases so that the maximum correlation between the image pixels can be maintained in the encrypted image. Additionally, a room is created inside the encrypted image for embedding the secret data. Therefore, the technique archives higher embedding capacity than the erstwhile state-of the art RDHEI techniques. In 2018, Li et al. [27] disclosed a novel RDHEI technique which makes use of combined block permutation and stream cipher for image encryption. Next, prediction errors in nearby pixels are exploited for data hiding. Thus, Li et al.'s technique improves the embedding rate by 0.5 bpp in comparison to the existing related RDHEI techniques, which signifies high embedding capacity.

In literature, it has been observed that a number of techniques [28–31] have been introduced for MSB prediction and then exploiting the MSB for data hiding. In 2018, one such work is proposed by Puteaux et al. [29] in which a simple but powerful high capacity RDHEI technique was discussed. The technique utilizes the correlation with neighboring pixel to predict the reference pixel value and uses a location binary map for marking the prediction errors. Next, the image is encrypted using stream cipher, and embedding of the secret message is done in the MSBs of the encrypted pixels with the help of location binary map. The location binary map helps the receiver in extracting the secret message and in the complete recovery of the original image. However, the MSB prediction technique can only embed up to one bit of the secret message into an encrypted pixel. To overcome this limitation, Puyang et al. [30] discuss a new RDHEI technique which can replace up to two MSBs of an encrypted pixel to significantly improve the embedding capacity without affecting the reversibility. To further improve the EC, Chen et al. [31] makes use of run length encoding (RLE) to compress the binary sequence of MSB data so that a room can be reserved in the image. For optimal compression, the image is first divided into blocks and

then the binary sequence of MSB data is created. Thus, the technique further boosts the embedding rate in comparison to [29,30] without compromising the reversibility.

Yi et al. [32] propounded a headway technique in the separable RDHEI domain using parametric binary tree labeling (PBTL). The host image is first partitioned into blocks of a certain size (either 2×2 or 3×3 pixels) and then calculates prediction errors in neighboring pixels within the block. Next, the host image is encrypted using a block-based encryption method. Based on the prediction errors and PBTL scheme, pixels of the encrypted are labeled which are exploited for embedding the secret information later on. The PBTL-RDHEI is extended by Su et al. [33] by combining the PBTL and absolute moment block truncation coding (AMBTC) for efficiently exploiting the correlation of the host image. Yin et al. [26] used the AMBTC technique for data embedding, although not in the RDHEI domain. Further, it has been observed that the AMBTC technique is a popular lossy compression technique which has been widely used in data hiding [14,36–38]. As far as working method of [33] is concerned, Su et al.'s technique first scrambles and then compresses the host image into triplets in a block-wise manner, where each triplet includes two quantization level (high & low) and a bitmap. Next, the triplets are encrypted in such a way that the correlation between the two quantization levels is retained. The retained correlation is then exploited to create a room for embedding the secret information. Therefore, a compressed marked-encrypted image in the form of AMBTC codes is obtained by the receiver. Since the bits required to represent the AMBTC coded image are lesser, the embedding capacity is also lower than that of Yi et al. [32]. It can be easily stated that Su et al.'s case is one of the earliest methods to utilize AMBTC in the RDHEI domain. In the next section, the parametric binary tree labeling scheme is reviewed in the context of the proposed work.

3. Parametric Binary Tree Labeling Scheme

This Parametric binary tree labelling (PBTL) scheme is first postulated by Yi et al. in 2018 for labelling the pixels of the plaintext image so that the auxiliary information can be efficiently embedded. The plaintext image pixels are represented by 8-bit depth; hence, the seven layers are generated in the parametric binary tree as shown in Figure 2. In the PBTL, each layer has 2^n nodes and each node is represented by n binary bits where $n \in \{1, 2, \ldots, 7\}$ is the layer number. The details regarding the total number of nodes in each layer of the seven-layered PTBL is provided in Table 1.

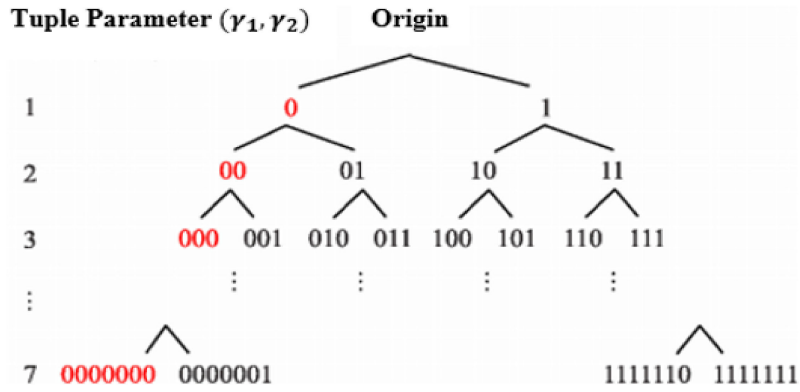


Figure 2. A 7-layered PBTL structure for 8-bit depth plaintext image.

Table 1. Number of nodes in a 7-layered PBTL.

Layer	1	2	3	4	5	6	7
Number of Nodes	2	4	8	16	32	64	128

The nodes at each layer are distributed into two different sets, namely S_1 and S_2, based on a tuple parameter (γ_1, γ_2). The first node of eachlayer is assigned into set S_1 and is labeled by n number of zero bits based on the value of γ_1, where n is the layer number. The number of nodes assigned to set S_2 are determined in accordance to the relation between the tuple parameter $\gamma_1 \wedge \gamma_2$ using Equation (1), defined as follows:

$$N_{\gamma_2} = \begin{cases} 2^{\gamma_2} - 1, \ \gamma_2 \leq \gamma_1 \\ (2^{\gamma_1} - 1) * (2^{\gamma_2 - \gamma_1}), \ \gamma_2 \leq \gamma_1 \end{cases} \quad (1)$$

where N_{γ_2} represents the total number of nodes in the set S_2 and the range of values of the tuple parameter is defined as $1 < \gamma_1, \gamma_2 < 7$. The specific details regarding the number of nodes in set S_2 for each layer of the PBTL based on the different values of $\gamma_1 \wedge \gamma_2$ are provided in Table 2. As per Table 2, if the value of tuple parameter $(\gamma_1, \gamma_2) is (2, 3)$, then number of nodes in set S_2 are 6 as per Equation (1), which are labeled as (111_2), (110_2), (101_2), (100_2), (011_2), and (010_2). Basically, the node labelling in the set S_2 is started from the right-most side to the left side of the PBTL (in accordance with Figure 2) until the N_{γ_2} is not labelled. Further, the first node in set S_1 will be labeled as (000_2).

Table 2. N_{γ_2}-based on $\gamma_1 \wedge \gamma_2$ for 7-layered PBTL.

γ_1	**Number of Nodes (N_{γ_2}) in S_2**						
	γ_2						
	1	**2**	**3**	**4**	**5**	**6**	**7**
1	1	2	4	8	16	32	64
2	1	3	6	12	24	48	96
3	1	3	7	14	28	56	112
4	1	3	7	15	30	60	120
5	1	3	7	15	31	62	122
6	1	3	7	15	31	63	126
7	1	3	7	15	31	63	127

4. The Proposed EPBTL-RDHEI Technique

The proposed High Capacity RDHEI technique is described in two phases, where the subscriber's data encryption and secret payload embedding is done in the first phase, and subscriber's data decryption and extraction of the secret payload is done in the second phase. Usually, subscribers upload their data which can be image, text, video, audio, etc., in encrypted form on the cloud server and a cloud administrator associates secret payload (which also includes auxiliary information) with the encrypted data. Although the proposed technique is equally applicable to all types of data, for simplicity, the plaintext images are considered as subscriber data. The proposed technique is a separable RDHEI technique which allows separate extraction of the hidden message and recovery of the original image at the receiving end. The RDHEI technique first uses an AMBTC based method for classifying the image regions and then uses a PBTL scheme for labelling the pixels. Figure 3 illustrates a framework of the proposed RDHEI technique.

The first phase of the proposed technique performed in five steps. In the first step, the subscriber's image is uniformly partitioned into non-overlapping blocks of a predetermined size and then each block is categorized into as smooth, moderately complex, and highly complex block based on the difference between high and low quantization levels which are calculated using the AMBTC method. In the second step, prediction errors and residual errors are generated based on the difference between pixel and corresponding block's mean. In the third step, the subscriber's image is encrypted using a stream encryption technique. In the fourth step, pixels of the encrypted image are categorized into regular pixels, irregular pixels, base pixels, and special pixels. Then, regular pixels and irregular

pixels are labeled according to the PBTL scheme and thereby labeled encrypted image is obtained. In the final (i.e., fifth) step, the secret payload embedding is done to get the marked encrypted image. The detailed working of each step is discussed in following subsections.

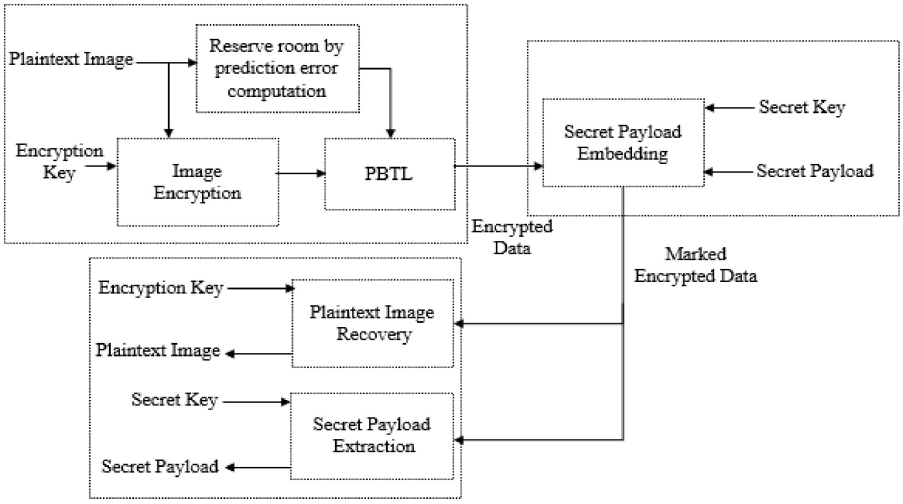


Figure 3. A framework of the proposed EPBTL-RDHEI technique.

4.1. Block Categorization (Step-1)

Initially, the subscriber-provided plaintext image $\mathbb{I}$ of size $N_1 \times N_2$ is partitioned into a number of non-overlapping blocks (P) of size $n_1 \times n_2$ pixels where $P = [N_1/n_1].[N_2/n_2]$. Therefore, each block (B_k) has $n_1 \times n_2$ pixels, such that, $\{x_1, x_2 \ldots . x_{n_1 \times n_2}\}$ if scanned in raster scan manner, where $k \in (1, 2 \ldots P)$. Now, mean (μ) of each block is computed using Equation (2):

$$\mu = \frac{1}{n_1 * n_2} \sum_{i=1}^{n1*n2} (x_i) \tag{2}$$

After computing mean (μ) of each block B_k, a bit-plane (b) of size $n_1 \times n_2$ is formed, where every pixel is represented by either bit '0' or '1' as follows.

$$b_i = \begin{cases} 0 \ if \ x_i < \mu, \\ 1 \qquad else. \end{cases} \tag{3}$$

As per the Equation (3), the pixel x_i of block B_k is represented by '0' if its value is less than mean (μ) of the block, otherwise represented by '1'. Thus, the bit-plane (b) for B_k has $n_1 \times n_2$ bits, meaning every pixel of the block is represented by 1 bit in the bit-plane. Next, two quantization levels, i.e., high quantization level q_0 and low quantization level q_1, are computed using Equations (4) and (5), respectively.

$$q_0 = \frac{1}{b_0} \sum_{x_i \geq \mu} x_i \tag{4}$$

$$q_1 = \frac{1}{b_1} \sum_{x_i < \mu} x_i \tag{5}$$

where b_0 and b_1 represent number of zeros and number of ones, respectively, in b. Thus, the high quantization level (q_0) is calculated by taking the mean of the pixels which have a value greater than the mean of the block. Similarly, the low quantization level (q_1) is computed by taking the mean of the pixels which have a value lower than the mean of the block. Thus, for each image block, a triplet $\{q_0, q_1, b\}$ is obtained.

Next, the image blocks are categorized into smooth, moderately complex, and highly complex block categories based on the difference between q_0 and q_1. More specifically, if the difference between q_0 and q_1 of block B is less than the first user-defined threshold T_s, then the block B is characterized as a smooth block type; if the difference between q_0 and q_1 is greater than or equal to the threshold T_s and less than a second user-defined threshold T_c, then the block B is characterized as moderately complex; and, otherwise, the block B is characterized as a highly complex block. In this step, the number of smooth blocks are counted as P_s, the number of moderately complex blocks are counted as P_m, and the number of high complex blocks are counted as P_h. A smooth block indicates that pixels in the block are high correlated and less distributed. A moderately complex block indicates that pixels in the block are less correlated and moderately distributed. A highly complex block indicates that pixels in the block are uncorrelated (or very little correlated) and highly distributed.

4.2. Computation of Prediction and Residual Errors (Step-2)

Once all the blocks of the plaintext image $\mathbb{I}$ are categorized, then prediction errors and residual errors are computed to reserve room for secret payload embedding as follows:

- If the block (B_k), where $k \in (1, 2 \ldots P)$ is a smooth block, then it indicates that pixels in the block B_k are high correlated and less distributed. Thus, each pixel in smooth block B_k can be best predicted by its mean value μ_k. Prediction error e_i is determined using the difference between the original pixel x_i and the mean value μ_k of the block (B_k) as per Equation (6). It is to be noted that the original pixel value (x_i) can be recovered by adding the μ_k to the corresponding prediction error (e_i) as per Equation (7), at the decoding side.

$$e_i = x_i - \mu_k \tag{6}$$

$$x_i = e_i + \mu_k \tag{7}$$

- If the block (B_k) is a moderately complex block, then it indicates that pixels in B_k are less correlated. In case of a moderately complex block, prediction errors (e_i) and residual errors (r_i) need to be calculated using Equations (8) and (9), respectively. It is to be noted that the original pixel value (x_i) can be recovered using Equation (10), at the decoder side.

$$e_i = \lfloor x_i/2 \rfloor - \lfloor \mu_k/2 \rfloor \tag{8}$$

$$r_i = mod(x_i/2) \tag{9}$$

$$x_i = e_i + r_i + 2 * \mu_k \tag{10}$$

- If the block (B_k) is a highly complex block, then it indicates that that pixels in B_k are uncorrelated (or very less correlated) and highly distributed. Thus, it requires a greater number of bits to represent the errors, and it is suggested that highly complex blocks are not used for secret payload embedding.

4.3. Image Encryption (Step-3)

After calculating the prediction errors and mean values for smooth blocks, and predication errors, residual errors, and mean values for moderately complex blocks, a stream encryption process is performed on the plaintext image $\mathbb{I}$. For this, a pseudo-random matrix $\mathbb{Z}$ of size $N_1 \times N_2$ is generated using an encryption key K_e. Each pixel $x_{i,j}$ of the image $\mathbb{I}$ is converted into an 8-bit binary sequence as $(b_8 b_7 b_6 b_5 b_4 b_3 b_2 b_1)$, where $b \in (0, 1)$ using Equation (11). Similarly, each element $z_{i,j}$ of pseudo-random matrix $\mathbb{R}$ is converted into 8-bit binary sequence using Equation (11).

$$x_{i,j}^k = \left\lfloor x_{i,j}/2^{k-1} \right\rfloor mod(2), k = (1, 2 \ldots 8) \tag{11}$$

Then, bitwise exclusive-or-operation is performed between plaintext image $\mathbb{I}$ and pseudo-random matrix $\mathbb{Z}$ using Equation (12) to encrypt the image.

$$xe_{i,j}^{k} = x_{i,j}^{k} \oplus z_{i,j}^{k} \tag{12}$$

Therefore, an encrypted 8-bit binary sequence for each pixel of the encrypted image is obtained. The binary sequence is converted into decimal form using Equation (13), to obtain an encrypted pixel which in turn helps in getting the encrypted image I^e.

$$xe_{i,j} = \sum_{k=1}^{8} xe_{i,j}^{k} \cdot 2^{k-1} \tag{13}$$

4.4. Pixel Grouping and Labelling Using PBTL (Step-4)

In this step, pixels of the encrypted image $\mathbb{I}^e$ are grouped and then labeled using PBTL structure based on tuple parameter (γ_1, γ_2) as described in Section 3. Firstly, the encrypted pixels of $\mathbb{I}^e$ are grouped into four sets which are special set (x^s), a base set (x^b), a regular set (x^r) and an irregular set (x^i). The special set (x^s) contains special pixels, the base set (x^b) contains base pixels, a regular set (x^r) contains regular pixels, and an irregular set (x^i) contains irregular pixels. Special set (x^s) includes the last pixel xe_{N_1,N_2} and the second last pixel $xe_{N_1-1,\ N_2-1}$ of the encrypted image $\mathbb{I}^e$. However, other locations can also be used in special set (x^s). One of the special pixels is used to indicate block size and other one indicates a tuple parameter (γ_1, γ_2). A base set (x^b) includes all the first encrypted pixel $xe_{1,1}$ of each block, where MSB bit-planes of the base pixels (the first encrypted pixel $xe_{1,1}$) are used to indicate block type that can be smooth, moderately complex, and highly complex. Table 3 illustrates that how the 7th & 8th MSB bit-planes of the encrypted pixel $xe_{1,1}$ is being used to identify block type.

Table 3. Block type identification based on 7th–8th bit-plane of $xe_{1,1}$ of each block B_k.

Bit-Plane of $xe_{1,1}$	**8th Bit-Plane**	**7th & 8th Bit-Planes**	**7th & 8th Bit-Planes**
Bit-Value	0	01	11
Block Type	Smooth block	Moderate complex block	High complex block

Pixels in the regular set (x^r) and irregular set (x^i) are determined in accordance to prediction errors $(e_{i,j})$ calculated using Equation (5) for smooth block and using Equation (7) for moderately complex block. If prediction errors $(e_{i,j})$ of block (B_k) of $\mathbb{I}$ meets the condition mentioned in Equation (13), then the pixel $(xe_{i,j})$ of the encrypted image $\mathbb{I}^e$ is grouped into the regular set (x^r), otherwise, it is grouped into the irregular set (x^i). In this step, the number of regular pixels and irregular pixels in smooth blocks are counted as x_s^r and x_s^i, respectively, and the number of regular pixels and irregular pixels in moderately complex blocks are counted as x_m^r and x_m^i, respectively.

$$\lceil -N_{\gamma_2}/2 \rceil \leq P_e \leq \lceil (N_{\gamma_2} - 1)/2 \rceil \tag{14}$$

After grouping the pixels into four sets, pixels of sets x^r and x^i are labeled using the PBTL scheme. Note that the pixels of highly complex blocks are not labeled as they do not participate in the data embedding process. In x^i, the γ_1-LSB bits of all the pixels are labeled as $(0...0_2)$ as per the PBTL scheme. In x^r, γ_2-LSB bits of all the pixels are labeled using different binary sequences represented by N_{γ_2} sub-categories as per the PBTL scheme. Thus processed, the labeled encrypted image I^l is obtained.

4.5. Secret Payload Embedding (Step-5)

After pixel grouping and pixel labelling process, secret payload embedding process is carried-out in the labeled encrypted image $\mathbb{I}^l$. This step outputs a marked encrypted image $\mathbb{I}^m$ by replacing the (8- γ_2) bits of pixels of x^r by secret payload. The secret payload (S_b) contains two type of data; one is encrypted auxiliary information (A_b) which is provided by cloud administrator and other is overhead (O_b) which is obtained during transforming encrypted image $\mathbb{I}^e$ into labeled encrypted image $\mathbb{I}^l$. To protect the auxiliary information (A_b) from unauthorized access, it is encrypted using a secret key k_s. The overhead (O_b) is required for lossless recovery of original image. The length of the overhead (O_b) is calculated as below:

(a) 16 bits to represent original pixel values of $xe_{N_1-1,\ N_2-1}$ and $xe_{N_1,\ N_2}$.
(b) Total bits (*equal to* $P_s + 2 * P_m + 2 * P_h$) of base set to store original bit values which are used in block types indication.
(c) Total bits (equal to $\gamma_2 * x_s^i + \gamma_2 * x_m^i$) of irregular set to store original replaced bit-planes.
(d) Total bits (equal to $8 * P_s$) to store mean value μ_k of each smooth block.
(e) Total bits (equal to $7 * P_m$) to store mean value $\mu_k/2$ of each complex block.

Now, auxiliary information (A_b) is obtained by reducing overhead (O_b) from the secret payload (S_b) using Equations (15)–(17), and effective embedding rate $ER_{\gamma_1,\ \gamma_2}$ is calculated using Equation (18).

$$A_b = S_b - O_b \tag{15}$$

$$S_b = (8 - \gamma_2) * x^{rs} + (7 - \gamma_2) * x^{rm} \tag{16}$$

$$O_b = (16 + \gamma_2 * x^i + 8 * P_s + 7 * P_m + P + P_m + P_h) \tag{17}$$

$$ER_{\gamma_1,\gamma_2} = \frac{A_b}{N_1 * N_2} = \frac{S_b - O_b}{N_1 * N_2} \tag{18}$$

The second phase of decryption of the original image and data recovery from the marked encrypted image has two steps. The first step is related to extraction of auxiliary information from the marked encrypted and second step is to restore original image from marked encrypted image. Both these steps are mutually exclusive.

4.6. Extraction of Auxiliary Information

At the receiver end, once the marked encrypted image $\mathbb{I}^m$ is received, the extraction process for auxiliary information is started. Initially, the pixels of x^s at locations xe_{N_1-1,N_2-1} and xe_{N_1,N_2} in $\mathbb{I}^m$ are utilized to determine tuple parameter $(\gamma_1,\ \gamma_2)$ and image block size. Then, $\mathbb{I}^m$ is divided into number of non-overlapping blocks as per the determined block size. After this, pixels of x^b in each image block are utilized to determine the type of the image block i.e. the smooth block, moderately complex block and highly complex block as per Table 3. Now, the tuple parameter $(\gamma_1,\ \gamma_2)$ is utilized to determine pixels x_s^r of regular set and pixels x_s^i of irregular set corresponding to smooth blocks, and pixels x_m^r of regular set and pixels x_m^i of irregular set corresponding to moderately complex blocks. Then, secret payload is extracted from (8- γ_2) bit-planes of regular pixels x_s^r of smooth blocks, and also from (8- γ_2) bit-planes of regular pixels x_m^r of moderately complex blocks. Further, residual errors are also extracted from regular pixels x_m^r of moderately complex blocks. Now, from the secret payload, encrypted auxiliary information is obtained by removing overhead as per Equations (15)–(17). Further, encrypted auxiliary information is decrypted by only authorized user which has secret key k_s.

4.7. Recovery of Plaintext Image

Now, recovery process for encrypted auxiliary information is discussed. The overhead information is utilized to recover plaintext image $\mathbb{I}$. Initially, 16 bits of overhead corresponding to $xe_{N_1-1,\ N_2-1}$ and $xe_{N_1,\ N_2}$. are restored on their predetermined position using x^s. Next, original bits (*equal to* $P_s + 2 * P_m + 2 * P_h$) are restored at locations of pixels of base set. Then, subsequent overhead bits equal to length $\gamma_2 * (x_s^i + x_m^i)$ are utilized to restore γ_2-bits of irregular pixels. After this, next overhead bits up to $8 * P_m$ are read to determine mean values of smooth blocks and subsequent overhead bits up to $7 * P_s$ are read to determine mean values of moderately complex blocks. Now, the authorized user which has encryption key K_e can generate the pseudo-random matrix $\mathbb{Z}$ and can obtain the decrypted image using Equations (12) and (13). Now, the user has only original pixel values of first two rows and two columns, which are utilized to recover other pixels of the image. The original pixel values of regular pixels of smooth blocks are obtained by their corresponding mean values stored as $8 * P_s$ overhead bits and prediction errors P_e calculated using Equation (7) from PBTL labeled pixels corresponding to $\gamma_2 - bits$. Similarly, to get original values of regular pixels of moderately complex blocks, mean values stored in next $7 * P_m$ overhead bits and residual errors P_r and prediction errors P_e calculated using Equation (10) from PBTL labeled pixels corresponding to $\gamma_2 - bits$. Now, remaining overhead bits are related to irregular pixels which are restored at their original location based on PBTL labeled pixels corresponding to $\gamma_1 - bits$. Thus, original image is retrieved in lossless manner.

Figure 4 shows an illustrative example of proposed RDHEI technique. The tuple parameter is taken as $(\gamma_1,\ \gamma_2) = (2, 3)$ and block size is taken as 4×4. Figure 4a shows a part of the baboon image of size 12×4. The input image is partitioned into non-overlapping blocks of size 4×4. Now, the blocks are categorized into smooth blocks, moderately complex blocks, and highly complex blocks using their corresponding quantization levels and mean values calculated by Equations (3)–(5), considering first threshold parameter $T_s = 16$ and second threshold parameter $T_c = 32$. The absolute difference between quantization levels $q_0 - q_1$ of the first block is 3, so it is a smooth block. The absolute difference between quantization levels $q_0 - q_1$ of the second block is 29, so it is a moderately complex block. The absolute difference between quantization levels $q_0 - q_1$ of the third block is 32, so it is a highly complex block. Highly complex blocks do not participate in secret payload embedding because pixels in the block are uncorrelated or correlated very little. Thus, only smooth blocks and moderately complex blocks are utilized for data embedding. The mean value of the first block is computed as 190 and the mean value of second block is computed as 103 using Equation (3). In Figure 4b,c, prediction errors are shown for each pixel of the block except for highly complex block. Based on the prediction errors, pixels of special set (x^s), base set (x^b), regular set (x^r) and irregular set (x^i) are determined as shown in Figure 4d. Now, input image is encrypted using stream encryption with encryption key K_e as shown in Figure 4e. Figure 4f shows an 8-bit binary representation of encrypted image. Using the tuple parameter (2,3), labelling bits for each pixel are determined PBTL structure. Pixels of the encrypted image corresponding to regular set (x^r) and irregular set (x^i) are labeled based on predictor errors. Figure 4h shows number of vacancies created in labeled regular pixels for embedding secret payload.

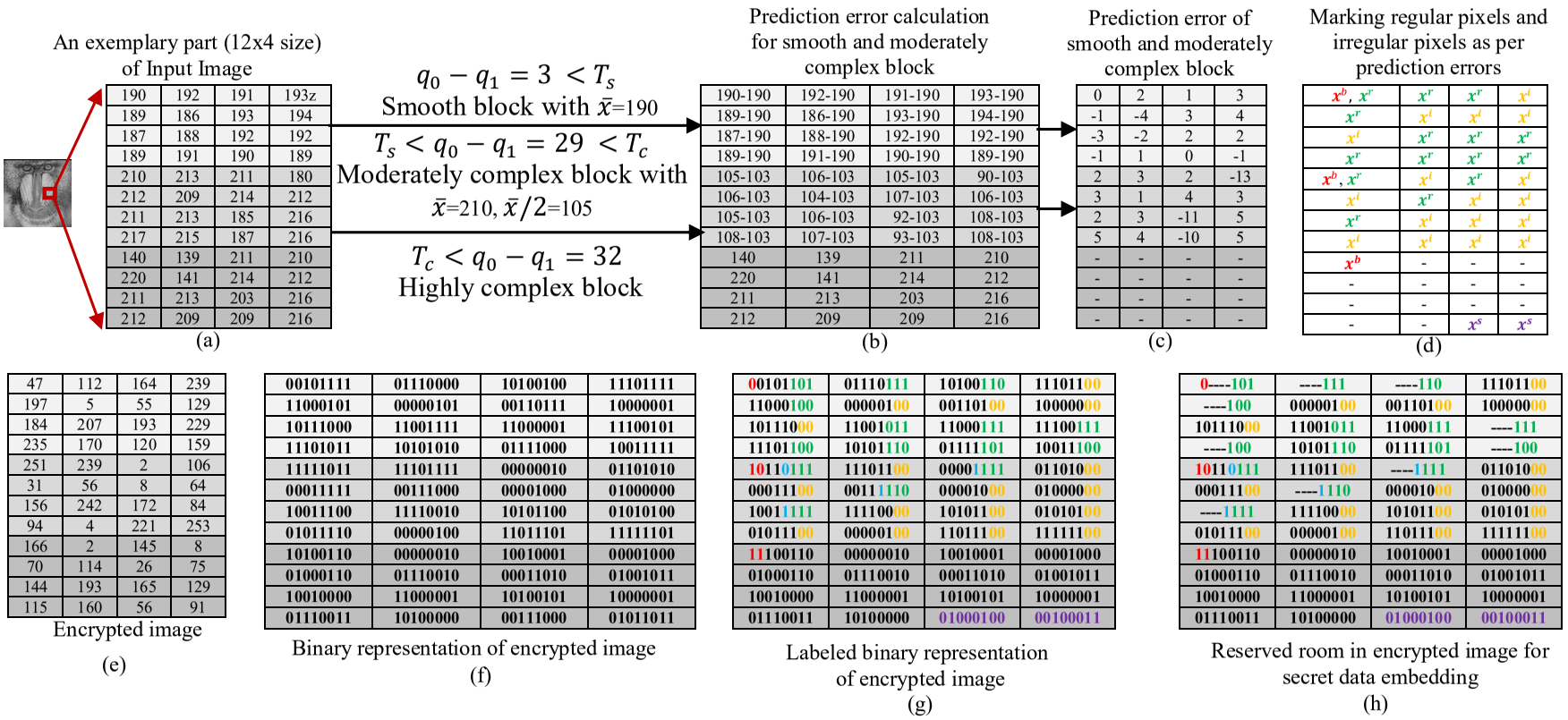


Figure 4. An illustrative example of the proposed scheme. (please resize to the format).

5. Experiment Results and Analysis

In this section, experiment results of the proposed RDHEI technique are discussed and compared with state-of-the-art techniques. To evaluate the performance of the proposed technique, we have taken four test images as shown in Figure 5. The test images are Lena, Airplane, Man, and Baboon, each one of size 512 × 512 with 8-bit depth gray values. The proposed technique has been evaluated using encryption performance and embedding rate. Encryption performance is estimated on two quality parameters, PSNR (peak signal-to-noise ratio) and SSIM (structural similarity), and one security parameter. Embedding rate (ER) basically represents embedding performance, which is measured in terms of bits per pixel (bpp).



(**a**) Lena (**b**) Airplane (**c**) Man (**d**) Baboon

Figure 5. Test images: (**a**) Lena, (**b**) Airplane, (**c**) Man, (**d**) Baboon.

5.1. Encryption Performance of Proposed Technique

In this subsection, encryption performance of the proposed technique is examined by computing pixel distributions of the marked encrypted images. Ideally, it should be uniform for the encrypted data as the encryption process completely destroys correlations present in original data to provide robustness. The proposed technique also provides uniform pixel distribution which is evident from Figure 6a–f, showing histograms of all test images, and Figure 6g–l, showing histograms of marked encrypted images for all the test images. The *X*-axis and *Y*-axis of the histograms represents number of pixels and intensity range (0–255), respectively.

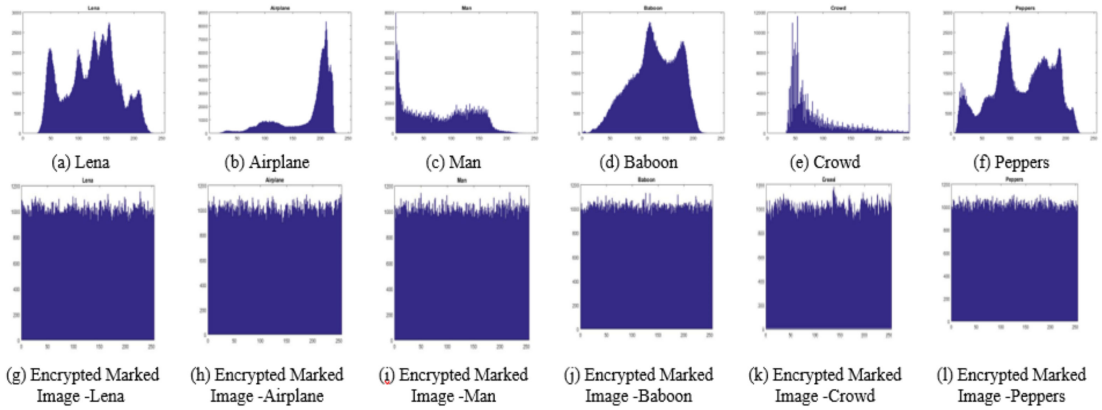


Figure 6. (**a**–**f**) histogram of test images, and (**g**–**l**) histograms of the encrypted marked images.

To further show the encryption performance of proposed technique, PSNR and SSIM values for marked encrypted images corresponding to the original test images are calculated and the results are provided in Table 4. The results are taken by considering tuple parameter (γ_1, γ_2) = (4,4) and block size of 4 × 4 pixels. It is to be noted that the experimental results on other tuple parameter values and block size are also similar. From Table 4, it can be clearly seen that PSNR of each encrypted marked image is very small and the SSIM value is also nearly 0, which indicates that encrypted marked image does not provide any information about the original image and secret payload.

Table 4. PSNR and SSIM values of encrypted marked image.

Test Image	PSNR (dB)	SSIM
Lena	9.2629	0.0203
Airplane	9.0388	0.0355
Man	8.3069	0.0260
Baboon	9.0285	0.0630
Crowd	9.1914	0.0116
Peppers	9.0116	0.0145

Thus, it is validated that the proposed technique provides good encryption performance. Figure 7a–f show the outcomes of different stages of image encryption and secret payload embedding for one of the test images. Figure 7a shows an original Lena image and Figure 7b shows an encrypted Lena image which is encrypted using the encrypted key K_e by employing a standard stream encryption algorithm.

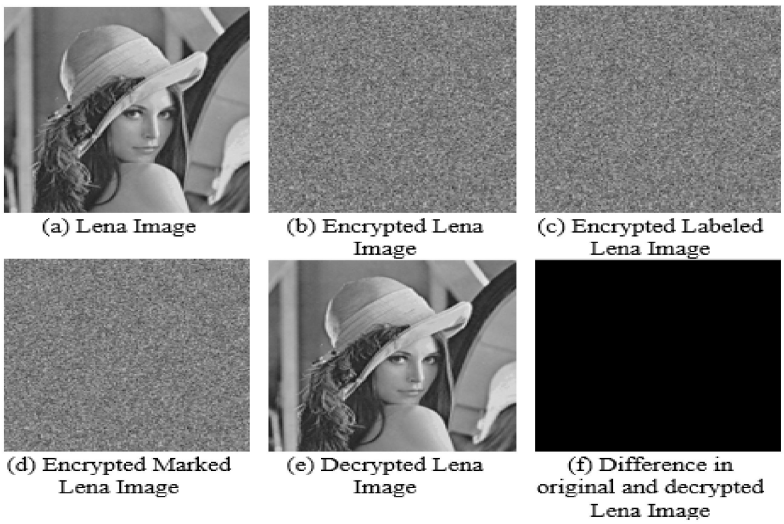


Figure 7. Encryption and decryption of Lena image.

In Figure 7c, an encrypted labeled Lena image is shown and Figure 7d shows a marked encrypted Lena image which includes a secret payload containing the encrypted auxiliary information, encrypted using secret key K_s. In Figure 7e, a decrypted Lena image is shown, which is similar to the original Lena image. In Figure 7f, the difference between Figure 7a,e is seen, showing a black image as all the pixel values are zero. Thus, Figure 7e depicts that the original image is fully recovered as it has PSNR $\rightarrow +\infty$ and SSIM = 1 with respect to the original image.

5.2. Embedding Performance of the Proposed Technique

In this subsection, embedding rate (ER) of the proposed technique is discussed in terms of bpp. The embedding rate has been shown for various values of tuple parameters (γ_1, γ_2 of the PBTL scheme and on different block sizes 4×4, 8×8, 16×16. Further, the thresholds are taken as $T_s = 5$, $T_c = 25$ to categorize image blocks as smooth, moderately complex, and highly complex blocks. Experimental results are provided in Tables 5–8 for test images. From the experimental results, it can be seen that embedding rate is increased, when tuple parameters are also increased up to a limit. Thereafter, embedding rate starts to decrease as tuple parameters are increased. This is because increasing in tuple parameters reduces vacancies for embedding secret data. Further, embedding rate is decreased for parameter γ_2 = 3, 4, 5 when block size is increased and the embedding rate is increased for

parameter γ_2 = 1, 2, 6, 7. The negative values of bpp depict that overhead bits are higher than or approximately equal to reserved room for secret data embedding.

Table 5. Embedding rates of Lena image at different values γ_1 and γ_2 for 4 × 4, 8 × 8, 16 × 16 blocks.

Lena Image, Block Size 4 × 4							
				γ_2			
γ_1	1	2	3	4	5	6	7
1	−0.3802	0.2735	1.2016	1.8779	1.5697	0.7723	−0.085
2	−1.1117	0.5411	1.8115	2.2203	1.6197	0.7736	−0.085
3	−1.8432	0.0459	2.018	2.2900	1.6251	0.7737	−0.085
4	−2.5747	−0.4493	1.8354	2.3141	1.6265	0.7738	−0.085
5	−3.3062	−0.9445	1.6528	2.289	1.6266	0.7737	−0.085
6	−4.0378	−1.4397	1.4701	2.2639	1.6258	0.7737	−0.085
7	−4.7693	−1.9349	1.2875	2.2388	1.6249	0.7737	−0.085
Block Size 8 × 8							
				γ_2			
γ_1	1	2	3	4	5	6	7
1	−0.0894	0.4042	1.1410	1.7434	1.5565	0.8958	0.1655
2	−0.7225	0.5738	1.6104	2.0712	1.6096	0.8983	0.1656
3	−1.3556	0.1291	1.7698	2.1362	1.6177	0.8987	0.1656
4	−1.9887	−0.3156	1.5864	2.1544	1.6197	0.8986	0.1656
5	−2.6218	−0.7603	1.403	2.1244	1.6199	0.8985	0.1656
6	−3.2548	−1.2051	1.2196	2.0944	1.618	0.8986	0.1656
7	−3.8879	−1.6498	1.0362	2.0644	1.6162	0.8985	0.1656
Block Size 16 × 16							
				γ_2			
γ_1	1	2	3	4	5	6	7
1	−0.085	0.2539	0.7632	1.2429	1.189	0.6916	0.1261
2	−0.5862	0.3104	1.0661	1.5345	1.2394	0.6947	0.1263
3	−1.0874	−0.0584	1.1793	1.6044	1.2473	0.6951	0.1264
4	−1.5886	−0.4273	1.0082	1.6205	1.2484	0.695	0.1264
5	−2.0898	−0.7961	0.8371	1.5898	1.2472	0.695	0.1264
6	−2.591	−1.1649	0.666	1.5591	1.2444	0.6949	0.1264
7	−3.0921	−1.5337	0.495	1.5284	1.2415	0.6946	0.1264

Table 6. Embedding rates of Airplane image at different values γ_1 and γ_2 for 4 × 4, 8 × 8, 16 × 16 blocks.

Airplane Image, Block Size 4 × 4							
				γ_2			
γ_1	1	2	3	4	5	6	7
1	0.0450	0.716	1.6555	2.0797	1.6411	0.8698	0.0542
2	−0.6023	1.2755	2.2124	2.3159	1.6763	0.8711	0.0542

Table 6. *Cont.*

Airplane Image, Block Size 4 × 4							
3	−1.2495	0.8933	2.3982	2.3613	1.6808	0.8712	0.0542
4	−1.8968	0.5111	2.2717	2.3693	1.6819	0.8712	0.0542
5	−2.544	0.129	2.1452	2.35	1.6822	0.8712	0.0542
6	−3.1913	−0.2532	2.0187	2.3307	1.6813	0.8712	0.0542
7	−3.8385	−0.6353	1.8922	2.3114	1.6805	0.8712	0.0542
Block Size 8 × 8							
				γ_2			
γ_1	1	2	3	4	5	6	7
1	0.2153	0.7296	1.4998	1.9414	1.6306	0.9878	0.2948
2	−0.3497	1.1507	1.9766	2.1751	1.6671	0.9899	0.2949
3	−0.9147	0.799	2.1473	2.2203	1.6736	0.9902	0.2949
4	−1.4797	0.4472	2.0201	2.2242	1.6749	0.9902	0.2949
5	−2.0447	0.0955	1.8929	2.2018	1.6739	0.9902	0.2949
6	−2.6098	−0.2563	1.7657	2.1794	1.672	0.9902	0.2949
7	−3.1748	−0.6081	1.6385	2.157	1.6701	0.9901	0.2949
Block Size 16 × 16							
				γ_2			
γ_1	1	2	3	4	5	6	7
1	0.1614	0.5260	1.1428	1.5400	1.3086	0.7920	0.2357
2	−0.3015	0.8476	1.5346	1.7483	1.3354	0.7946	0.2359
3	−0.7643	0.5505	1.6901	1.7828	1.3396	0.7947	0.236
4	−1.2271	0.2534	1.5821	1.7851	1.3406	0.7947	0.2361
5	−1.6899	−0.0437	1.4741	1.7664	1.3384	0.7943	0.2361
6	−2.1528	−0.3408	1.3662	1.7478	1.3357	0.794	0.2361
7	−2.6156	−0.6379	1.2582	1.7292	1.333	0.7936	0.2361

Table 7. Embedding rates of Man image at different values γ_1 and γ_2 for 4 × 4, 8 × 8, 16 × 16 blocks.

Man Image, Block Size 4 × 4							
				γ_2			
γ_1	1	2	3	4	5	6	7
1	−0.4069	0.1733	0.9315	1.4777	1.2980	0.5978	−0.2026
2	−1.0968	0.3123	1.3616	1.8139	1.3764	0.6004	−0.2026
3	−1.7867	−0.1722	1.4769	1.8907	1.3885	0.6007	−0.2026
4	−2.4766	−0.6567	1.2620	1.9133	1.3912	0.6007	−0.2026
5	−3.1665	−1.1412	1.0471	1.8714	1.3919	0.6007	−0.2026
6	−3.8565	−1.6256	0.8321	1.8295	1.3902	0.6007	−0.2026
7	−4.5464	−2.1101	0.6172	1.7877	1.3884	0.6007	−0.2026

Table 7. *Cont.*

				Block Size 8 × 8			
				γ_2			
γ_1	1	2	3	4	5	6	7
1	−0.1553	0.2655	0.8052	1.2536	1.1992	0.6775	0.0286
2	−0.7307	0.2760	1.0819	1.5326	1.2917	0.6838	0.0287
3	−1.3061	−0.1497	1.1359	1.6029	1.3090	0.6842	0.0287
4	−1.8815	−0.5755	0.9212	1.6142	1.3120	0.6843	0.0287
5	−2.4569	−1.0012	0.7066	1.5595	1.3125	0.6843	0.0287
6	−3.0323	−1.4270	0.4919	1.5048	1.3084	0.6843	0.0287
7	−3.6077	−1.8527	0.2773	1.4502	1.3042	0.6841	0.0287
				Block Size 16 × 16			
				γ_2			
γ_1	1	2	3	4	5	6	7
1	−0.1489	0.1368	0.4943	0.8332	0.8754	0.5204	0.0399
2	−0.5902	0.0772	0.6419	1.0590	0.9603	0.5275	0.0402
3	−1.0315	−0.2643	0.6461	1.1145	0.9773	0.5288	0.0402
4	−1.4728	−0.6059	0.4532	1.118	0.9811	0.5290	0.0403
5	−1.9141	−0.9474	0.2602	1.0616	0.9795	0.5289	0.0403
6	−2.3554	−1.2890	0.0673	1.0052	0.9734	0.5287	0.0403
7	−2.7967	−1.6305	−0.1256	0.9488	0.9672	0.5283	0.0403

Table 8. Embedding rates of Baboon at different values γ_1 and γ_2 for 4 × 4, 8 × 8, 16 × 16 blocks.

				Baboon Image, Block Size 4 × 4			
				γ_2			
γ_1	1	2	3	4	5	6	7
1	−0.5311	−0.3264	−0.0006	0.4225	0.6605	0.3702	−0.1360
2	−1.0109	−0.4916	0.0785	0.7055	0.8125	0.3801	−0.1360
3	−1.4906	−0.8955	0.0214	0.786	0.8445	0.3810	−0.1360
4	−1.9704	−1.2994	−0.2418	0.7943	0.8536	0.3812	−0.1360
5	−2.4502	−1.7033	−0.505	0.7081	0.8552	0.3813	−0.1360
6	−2.9299	−2.1072	−0.7682	0.6218	0.8490	0.3814	−0.1360
7	−3.4097	−2.5111	−1.0314	0.5356	0.8428	0.3813	−0.1360
				Block Size 8 × 8			
				γ_2			
γ_1	1	2	3	4	5	6	7
1	−0.2529	−0.1006	0.1333	0.4396	0.6186	0.4100	0.0215
2	−0.6273	−0.2436	0.1757	0.6395	0.7409	0.4217	0.0216
3	−1.0017	−0.5607	0.1120	0.6929	0.7663	0.4232	0.0216
4	−1.3760	−0.8779	−0.0999	0.6905	0.7723	0.4236	0.0216
5	−1.7504	−1.1951	−0.3117	0.6151	0.7712	0.4237	0.0216

Table 8. *Cont.*

6	−2.1248	−1.5122	−0.5236	0.5397	0.7633	0.4238	0.0216
7	−2.4992	−1.8294	−0.7355	0.4642	0.7553	0.4236	0.0216
				Block Size 16 × 16			
				γ_2			
γ_1	1	2	3	4	5	6	7
1	−0.1643	−0.052	0.1171	0.3358	0.4629	0.3089	0.0168
2	−0.4502	−0.1665	0.1343	0.4706	0.5529	0.3198	0.0174
3	−0.7361	−0.4093	0.0777	0.5044	0.5715	0.3212	0.0174
4	−1.0220	−0.6522	−0.0874	0.5004	0.5745	0.3214	0.0174
5	−1.3079	−0.8951	−0.2525	0.4383	0.5720	0.3213	0.0174
6	−1.5937	−1.1379	−0.4176	0.3761	0.5634	0.3210	0.0174
7	−1.8796	−1.3808	−0.5827	0.3140	0.5548	0.3204	0.0174

5.3. Comparison of Results of the Proposed Technique with State of Art Techniques

In this subsection, experimental results of the proposed techniques are compared with related techniques including those of Yi et al. [32], Su et al. [33], Li et al. [27], Puteaux et al. [29], Puyang et al. [30], Chen et al. [31]. Since PBTL was first introduced in the RDHEI domain by Yi et al., the same [24] has been primarily considered for comparison. Su et al.'s [33] technique demonstrates application of AMBTC concept along with PBTL. However, it is a lossy technique where a decrypted image is obtained in compressed form. Since the proposed technique also considers some aspects of AMBTC in block categorization, Su et al.'s [33] technique has also been used to comparatively evaluate the performance. However, the proposed technique is a lossless and fully reversible technique in which both secret payload and original image are retrieved in undistorted form. Additionally, the performance of the proposed technique is evaluated against some of the early high capacity RDHEI techniques, e.g., Li et al. [27], Puteaux et al. [29], Puyang et al. [30], and Chen et al. [31]. Table 9 shows the encryption method of the proposed technique and state-of-art techniques. In comparison to this, Yi et al. [32] retains some pixel correlations in blocks which makes encryption week.

Table 9. Comparison of encryption performance with state-of-art techniques.

Technique	Encryption Method	Preserving Pixel Redundancy
Yi et al. [32]	Block Permutation and block modulation	Yes
Su et al. [33]	Block Scrambling and Stream Encryption	Yes
Li et al. [27]	Block Permutation and Stream Encryption	No
Puteaux et al. [29]	Stream Encryption	No
Puyang et al. [30]	Stream Encryption	No
Chen et al. [31]	Stream Encryption	No
Proposed RDHEI	Stream Encryption	No

For optimal embedding performance, $\gamma_1 = 4$ & $\gamma_2 = 4$ and a block size of 4×4 pixels is considered for the proposed RDHEI technique. The embedding performance is compared with some of the high-capacity state-of-the-art techniques that provide highest embedding

rate. Since Yi et al.'s [32] and Chen et al. [33] have optimal performance at $\gamma_1 = 2$ & $\gamma_2 = 5$ with block size of 3×3 pixels and block size of 4×4 pixels, respectively.

Figure 8 shows the maximal embedding rates of the proposed technique as well as existing RDHEI techniques for all the test images. It clearly shows that proposed technique has highest embedding rate than other techniques. It is also evident from the figure that the proposed scheme provides higher embedding capacity with smooth images like Lena, Airplane, whereas it has lower embedding capacity with complex images such as Baboon.

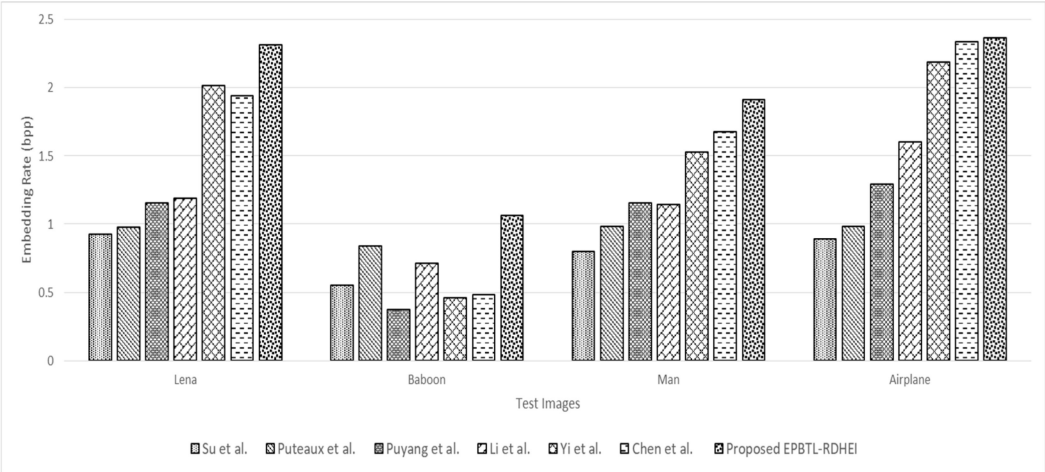


Figure 8. Comparison of embedding rate of the proposed and state-of-art techniques.

6. Conclusions

In this paper, an intra-block correlation based high capacity RDHEI technique using a parametric binary tree labelling scheme has been proposed. In the proposed technique, the cover image was divided into blocks and then the blocks were categorized according to the prevalent correlation through a symmetric or asymmetric process. Next, an adaptive method for reserving the room inside the image was applied, based on the block categories so that a large amount of secret data could be embedded. Further, the proposed RDHEI technique used a stream cipher for encrypting the image contents. Experimental results showed that the proposed RDHEI technique provided the highest embedding rate in comparison to all the aforementioned state-of-the art techniques. Additionally, the proposed method provided a good level of security in encryption process to protect the privacy of the original plaintext image. In future work, compression methods can be explored to further condense the size of prediction errors, and also an improvised AMBTC method may be designed to predict the pixel values more accurately.

Author Contributions: Conceptualization, A.K.R. and N.K.; Methodology, A.K.R. and N.K.; Visualization, R.K.; Writing—A.K.R., N.K.; Writing—review and editing, R.K., H.O., S.C. and K.-H.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Brain Pool program funded by the Ministry of Science and ICT through the National Research Foundation of Korea (2019H1D3A1A01101687) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2018R1D1A1A09081842 and 2021R1I1A3049788).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

Symbols	**Meaning**
A_b	auxiliary information
μ	Mean
B_k	Block
b_i	Logic bit value
K_e	Encryption Key
K_s	Secret Key
n	Number of Layers
γ_1, γ_2	Tuple Parameter
S_1, S_2	Two different sets
N_{γ_2}	Total number of nodes in the set S_2
$\mathbb{I}$	plaintext image
I^e	Encrypted image
$N_1 \times N_2$	Image Size
$n_1 \times n_2$	Block Size
O_b	Overhead
p	Number of non-overlapping blocks
q_0, q_1	Quantization values
T_s, T_c	Threshold values
x_i	i-th pixel
x^s	Special set
x^b	Base set
x^r	Regular set
x^i	Irregular set
Abbreviation	**Meaning**
AES	Advanced Encryption Standard
AMBTC	Absolute Moment Block Truncation Coding
BPP	Bit-Per-Pixel
EC	Embedding Capacity
ER	Embedding Rate
LSB	Least Significant Bits
MSB	Maximum Significant Bits
PBTL	Parametric Binary Tree Labeling Scheme
PE	Prediction Error
PSNR	Peak Signal-To-Noise Ratio
RDH	Reversible Data Hiding
RDHEI	Reversible Data Hiding in Encrypted Images
RLE	Run Length Coding
RRBE	Reserving Room Before Encryption
SSIM	Structural Similarity Index
VRAE	Vacating Room After Encryption
VRBE	Vacating Room by Encryption

References

1. Pahl, C.; Brogi, A.; Soldani, J.; Jamshidi, P. A State-of-the-Art Review. *IEEE Trans. Cloud Comput.* **2019**, *7*, 677–692. [CrossRef]
2. Chandramouli, R.; Iorga, M.; Chokhani, S. Cryptographic key management issues and challenges in cloud services. *Secure Cloud Comput.* **2014**, 1–30. [CrossRef]
3. Kumar, R.; Chand, S. A reversible high capacity data hiding scheme using pixel value adjusting feature. *Multimed. Tools Appl.* **2016**, *75*, 241–259. [CrossRef]
4. Lee, C.; Shen, J.; Wu, Y.; Agrawal, S. PVO-based reversible data hiding exploiting two-layer embedding for enhancing image fidelity. *Symmetry* **2020**, *12*, 1164. [CrossRef]
5. Khan, S.; Khan, K.; Arif, A.; Hassaballah, M.; Ali, J.; Ta, Q.; Yu, L. A modulo function-based robust asymmetric variable data hiding using DCT. *Symmetry* **2020**, *12*, 1659. [CrossRef]
6. Kim, S.; Cho, N.I. Hierarchical Prediction and Context Adaptive Coding for Lossless Color Image Compression. *IEEE Trans. Image Process.* **2014**, *23*, 445–449. [CrossRef] [PubMed]

7. Alattar, A.M. Reversible Watermark Using the Difference Expansion of a Generalized Integer Transform. *IEEE Trans. Image Process.* **2004**, *13*, 1147–1156. [CrossRef] [PubMed]
8. Kumar, R.; Jung, K.-H. Enhanced pairwise IPVO-based reversible data hiding scheme using rhombus context. *Inf. Sci.* **2020**, *536*, 101–119. [CrossRef]
9. Wu, H.; Li, X.; Zhao, Y.; Ni, R. Improved reversible data hiding based on PVO and adaptive pairwise embedding. *J. Real-Time Image Process.* **2019**, *16*, 685–695. [CrossRef]
10. Ou, B.; Li, X.; Wang, J. High-fidelity reversible data hiding based on pixel-value-ordering and pairwise prediction-error expansion. *J. Vis. Commun. Image Represent.* **2016**, *39*, 12–23. [CrossRef]
11. Gao, G.; Tong, S.; Xia, Z.; Wu, B.; Xu, L.; Zhao, Z. Reversible data hiding with automatic contrast enhancement for medical images. *Signal Process.* **2021**, *178*, 107817. [CrossRef]
12. Kumar, R.; Jung, K.-H. Robust reversible data hiding scheme based on two-layer embedding strategy. *Inf. Sci.* **2020**, *512*, 96–107. [CrossRef]
13. Puech, W.; Chaumont, M.; Strauss, O. A reversible data hiding method for encrypted images. In Proceedings of the SPIE, San Jose, CA, USA, 18 March 2008; pp. 68191E-1–68191E-9.
14. Bhardwa, J.R.; Aggarwal, A. An improved block based joint reversible data hiding in encrypted images by symmetric cryptosystem. *Pattern Recognit. Lett.* **2020**, *139*, 60–68. [CrossRef]
15. Hong, W.; Chen, T.-S.; Wu, H.-Y. An Improved Reversible Data Hiding in Encrypted Images Using Side Match. *IEEE Signal Process. Lett.* **2012**, *19*, 199–202. [CrossRef]
16. Zhou, J.; Sun, W.; Dong, L.; Liu, X.; Au, O.C.; Tang, Y.Y. Secure Reversible Image Data Hiding Over Encrypted Domain via Key Modulation. *IEEE Trans. Circuits Syst. Video Technol.* **2016**, *26*, 441–452. [CrossRef]
17. Wu, X.; Sun, W. High-capacity reversible data hiding in encrypted images by prediction error. *Signal Process.* **2014**, *104*, 387–400. [CrossRef]
18. Zhang, W.; Ma, K.; Yu, N. Reversibility improved data hiding in encrypted images. *Signal Process.* **2014**, *94*, 118–127. [CrossRef]
19. Zhang, X. Separable Reversible Data Hiding in Encrypted Image. *IEEE Trans. Inf. Forensics Secur.* **2012**, *7*, 826–832. [CrossRef]
20. Kaur, G.; Singh, S.; Rani, R.; Kumar, R.; Malik, A. High-quality reversible data hiding scheme using sorting and enhanced pairwise PEE. *IET Image Process.* **2021**, 1–15. [CrossRef]
21. Dragoi, I.C.; Coanda, H.-G.; Coltuc, D. Improved reversible data hiding in encrypted images based on reserving room after encryption and pixel prediction. In Proceedings of the 25th European Signal Processing Conference (EUSIPCO), Kos Island, Greece, 28 August–2 September 2017; pp. 2186–2190.
22. Dragoi, I.C.; Coltuc, D. Reversible Data Hiding in Encrypted Color Images Based on Vacating Room After Encryption and Pixel Prediction. In Proceedings of the 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 1673–1677.
23. Ma, K.; Zhang, W.; Zhao, X.; Yu, N.; Li, F. Reversible Data Hiding in Encrypted Images by Reserving Room Before Encryption. *IEEE Trans. Inf. Forensics Secur.* **2013**, *8*, 553–562. [CrossRef]
24. Xu, D.; Wang, R. Separable and error-free reversible data hiding in encrypted images. *Signal Process.* **2016**, *123*, 9–21. [CrossRef]
25. Mathew, T.; Wilscy, M. Reversible data hiding in encrypted images by active block exchange and room reservation. In Proceedings of the 2014 International Conference on Contemporary Computing and Informatics (IC3I), Mysore, India, 27–29 November 2014; pp. 839–844.
26. Cao, X.; Du, L.; Wei, X.; Meng, D.; Guo, X. High Capacity Reversible Data Hiding in Encrypted Images by Patch-Level Sparse Representation. *IEEE Trans. Cybern.* **2016**, *46*, 1132–1143. [CrossRef] [PubMed]
27. Li, Q.; Yan, B.; Hui, L.; Chen, N. Separable reversible data hiding in encrypted images with improved security and capacity. *Multi. Tools Apps.* **2018**, *77*, 30749–30768. [CrossRef]
28. Hang, Y.; Wei, T.; Menghan, G.; Shoushun, C. A two-step prediction ADC architecture for integrated low power image sensors. *IEEE Trans. Circuits Syst. I* **2017**, *64*, 50–60.
29. Puteaux, P.; Puech, W. An Efficient MSB Prediction-Based Method for High-Capacity Reversible Data Hiding in Encrypted Images. *IEEE Trans. Inf. Forensics Secur.* **2018**, *13*, 1670–1681. [CrossRef]
30. Puyang, Y.; Yin, Z.; Qian, Z. Reversible Data Hiding in Encrypted Images with Two-MSB Prediction. In Proceedings of the 2018 IEEE International Workshop on Information Forensics and Security (WIFS), Hong Kong, China, 11–13 December 2018; pp. 1–7.
31. Kaimeng, C.; Chang, C.C. High-capacity Reversible Data Hiding in Encrypted Images Based on Extended Run-Length Coding and Block-based MSB Plane Rearrangement. *J. Vis. Commun. Image Represent.* **2019**, *58*, 334–344.
32. Yi, Y.; Zhou, Y. Separable and Reversible Data Hiding in Encrypted Images Using Parametric Binary Tree Labeling. *IEEE Trans. Multimed.* **2019**, *21*, 51–64. [CrossRef]
33. Su, G.-D.; Chang, C.-C.; Lin, C.-C. A High Capacity Reversible Data Hiding in Encrypted AMBTC-Compressed Images. *IEEE Access* **2020**, *8*, 26984–27000. [CrossRef]
34. Yin, Z.; Niu, X.; Zhang, X.; Tang, J.; Luo, B. Reversible data hiding in encrypted AMBTC images. *Multimed. Tools Appl.* **2018**, *77*, 18067–18083. [CrossRef]
35. Yin, Z.; Wang, H.; Zhao, H.; Luo, B.; Zhang, X. Complete separable reversible data hiding in encrypted image. In Proceedings of the 1st International Conference on Cloud Computing and Security, Nanjing, China, 13–15 August 2015; pp. 101–110.

36. Wang, R.; Wu, G.; Wang, Q.; Yuan, L.; Zhang, Z.; Miao, G. Reversible Data Hiding in Encrypted Images Using Median Edge Detector and Two's Complement. *Symmetry* **2021**, *13*, 921. [CrossRef]
37. Kumar, R.; Kim, D.-S.; Jung, K.-H. Enhanced AMBTC based data hiding method using hamming distance and pixel value differencing. *J. Inf. Secur. Appl.* **2019**, *47*, 94–103. [CrossRef]
38. Kim, S. Reversible Data-Hiding Systems with Modified Fluctuation Functions and Reed-Solomon Codes for Encrypted Image Recovery. *Symmetry* **2017**, *9*, 61. [CrossRef]

2021
8(2), 59-66

Chemical Biology
LETTERS

Isolation, Synthesis, and Medicinal applications of Heparin

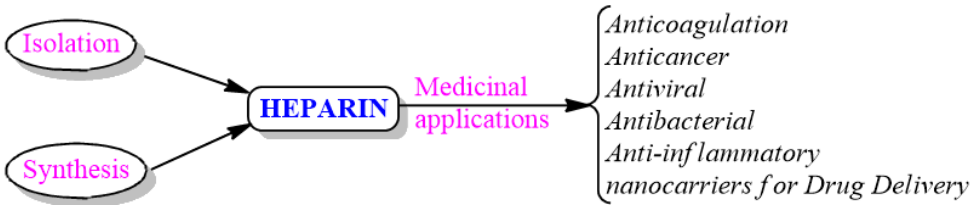
Saikrushna Jena, Ram Singh*

Department of Applied Chemistry, Delhi Technological University, Delhi-110 042, India

Submitted on: 25-Apr-2021, Accepted and Published on: 01-June-2021

Mini-Review Article

ABSTRACT



Heparin is a highly sulphated and most negatively charged natural biopolymer belonging to the glycosaminoglycan (GAG) family. This is about a 100-year-old anticoagulant drug. This is equally important for non-anticoagulant diseases also and is the reason for the recent burgeoning of interest in the molecule. Heparin has been isolated from both animal and non-animal sources; however, porcine mucosa remains the FDA-approved source for heparin. For the synthesis, chemical, chemoenzymatic, and biotechnological approaches have been studied. In recent times, the focus is more on synthesizing LMWH, ULMWH, and bioengineered heparins. This article is a compilation about the isolation, synthesis, and medicinal applications of heparin.

Keywords: Heparin, glycosaminoglycan, biopolymer, LMWH, ULMWH, anticoagulant

INTRODUCTION

Heparin (Figure 1) is a natural biopolymer belonging to the glycosaminoglycan (GAG) family.[1-3] This is a highly sulphated and most negatively charged polysaccharide. The heparin polysaccharide has disaccharide repeating units, where the disaccharide is 1,4-glycosidically linked and consists of uronic acid residues and D-glucosamine (GlcN). The uronic acid residue is either β-D-glucuronic acid or α-L-iduronic acid (IdoA).[4] The GlcN is further substituted with an N-acetyl (GlcNAc) or an N-sulpho (GlcNS) group, which can also be O-sulphated at the 6-position.[1] This molecule is heterogeneous in nature due to structural variations, which help to interact with different proteins leading to various biological activities. Some of the important properties of heparin are summarized in Figure 2.[5,6] Some of the important brands of heparin available with their generic name as enoxaparin, dalteparin, tinzaparin, heparin flush, danaparoid.

Many researchers have described the discovery and development of heparin. The three persons, William Henry Howell, Luther Emmett Holt (Jr), Jay McLean, names were involved in the discovery of this molecule in 1916-1918.[7,8] After about 14-15 years of work, this biopolymer was considered as an effective anticoagulant drug.[9] Further, medicinal investigation proved that heparin is equally potential as a pharmacological molecule for various non-anticoagulant diseases also.[6,10-12]

Being a biopolymer, heparin is biosynthesized by the concerted action of about 22 enzymes which reflects its complexity towards synthesis. This is biosynthesized in the endoplasmic reticulum (ER) and Golgi.[13-16] This molecule is isolated from porcine intestinal mucosa or bovine sources.[17,18] The isolation of this pharmaceutic is very tedious, with a low yield of about 180-260 mg per healthy animal. One of the major limitations of its production is that this molecule is mainly obtained from a single animal species. A lot of research has been done towards finding the other animal and non-animal sources with very little or no success. In non-animal source finding research, chemical and chemoenzymatic synthesis have been tried with very limited success. A most successful example of chemical synthesis is the chemical synthesis of Arixtra. The synthesis of this molecule takes place in about 50 steps with an overall yield of ~0.1%.[19,20] However, this is also not purely chemical synthesis.

"According to government data, the annual market for heparin is about Rs 50 crore for almost 63 lakh units and Rs 504 crores for enoxaparin, which is for about 150 lakh units. These estimates are of the product imported into India, which is about 2,500-3,000 kilos for Heparin from which about 1,000 Kgs of

*Corresponding Author: Dr. Ram Singh
Department of Applied Chemistry, Delhi Technological University, Delhi-110 042, India
Tel: NA
Email: ramsingh@dtu.ac.in

URN:NBN:sciencein.cbl.2021.v8.248 ISSN: 2347–9825
© ScienceIn Publishing http://pubs.thesciencein.org/cbl

enoxaparin is processed in India. This doesn't account for the product that is imported directly as the finished product.[21-23]"

This article is a compilation about the isolation, synthesis and medicinal applications of this biomolecule by taking the representative examples.

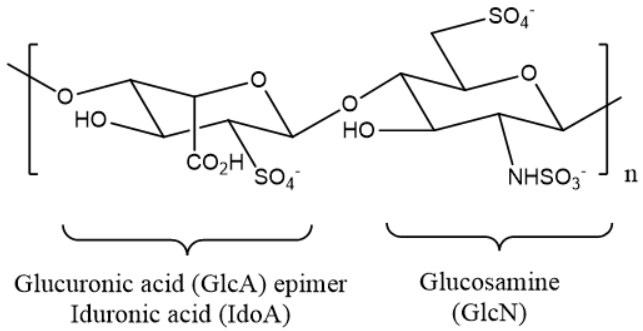


Figure 1. Maximum disaccharides in heparin

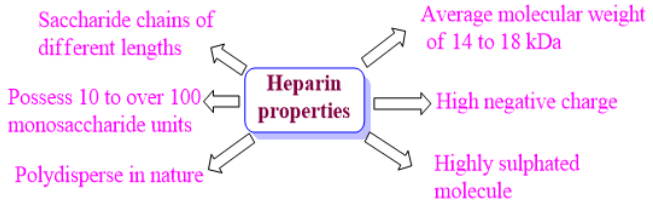


Figure 2. Important properties of heparin

ISOLATION AND SYNTHESIS OF HEPARIN

Isolation of heparin

Any animal used as the source for any drug isolation has to follow strict ethical guidelines. For heparin isolation also, the health and medication of slaughtered animals have to meet certain requirements.[24] There is full traceability from farm to pharma is maintained with proper regulations.[24-26] The earlier heparin isolation was done with canine or bovine livers; later, bovine lungs and mucosa were used (Figure 3).[18] However, due to the outbreak of mad cow disease (Bovine Spongiform Encephalopathy (BSE))[27], the use of bovine materials for heparin isolation decreased significantly.[28] The porcine mucosa remains the FDA-approved source for heparin; however, FDA encourages reintroduction of bovine-source also.[29] There is a strong requirement for enhanced knowledge about bovine disease and the heparin purification process to reduce the risk associated. Studies have shown that the bovine heparin is less active than the porcine heparin; this may be due to their structural differences between the two.[30]

The sheep (ovine) intestines have also been studied as a source to isolate pharmaceutical heparin.[28] Ovine heparin showed better disaccharide composition than the bovine heparin.[31,32] Further, the use of ovine is also free from any religious sentiments. However, the disease like scrapie in sheep found to have some concern, although this is not transmissible to humans.

Many mammalian sources have been studied to isolate heparin.[33,34] The human tissues were used as a natural source of heparin.[35,36] The frozen hemangioma tissue gave human heparin in the amount 649 µg/g of tissue.[35] Warda et al. isolated heparin from one-humped camel in 400 mg/kg amount from an adult camel.[37] This method and source gave better results in comparison to the source porcine intestine (250 mg/kg).

Among the non-mammalian sources, poultry, molluscs, fishes etc. have been explored to isolate heparin. Chicken intestines produced heparin with a lower degree of sulphation.[30] Heparin was also isolated from the turkey intestine but showed very poor activity.[38] Till today, to the best of my knowledge, there is no industry using poultry by-products to isolate heparin. In another study, tuna (saltwater fish) skins have been utilized to isolate heparin with the help of anion exchange resin.[39] Through steaming, the skin was obtained from the freshly caught tuna fish. The obtained heparin showed less potency than the commercial product.[39] Heparin was also isolated and purified from gills and intestines of salmon (*Salmo salar*) whose activity showed similarity with the activity of low-molecular-weight heparin (LMWH).[40] The heads of shrimp (*Penaeus brasiliensis*) produced LMWH with a yield of 32 mg/kg.[41] The isolation from clams (*Tapes phylippinarum*) gave heparin with a yield of ~2.1 g/kg dry tissue.[42] The commercial isolation/production from the discussed sources (Figure 3) is still a matter of research.

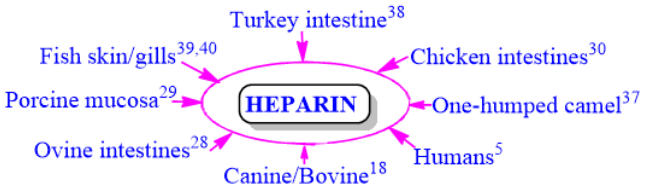


Figure 3. Natural sources of heparin

Synthesis of heparin

The synthesis of heparin is very challenging, and still a lot of work is going on towards this. Heparin of various polysaccharide chains has been synthesized/semi-synthesized.[43-46] As per the length and molecular weight, heparin is grouped as unfractionated heparin (UFH), low molecular weight heparin (LMWH), and ultra-low molecular weight heparin (ULMWH). Chemical synthesis and chemoenzymatic synthesis have been tried with limited success. In both the methods, oligosaccharide building method has been explored. The methods are multi-steps, taking place in more than 50-60 steps leading to very-very low yields. For example, a heparin, fondaparinux, was synthesized in about 60 steps with an overall yield of only 0.1%.[47,48] Fondaparinux is a pentasaccharide sequence having a similar binding region as heparin. The challenge associated with heparin synthesis are many like repetitive protection deprotection steps, couplings, functional group activation, intensive purification, and so on.[1,49]

The chemoenzymatic synthesis takes the help of polymerases which help in backbone building. The backbones have further been modified with enzymes like C5 epimerase and sulfotransferases.[50] The backbone of the chemoenzymatic method is the process involved in its biosynthesis; rather, one can say it is mimicking the biosynthesis of heparin.[51-53] Xu et al.

synthesized two ULMW heparins through a chemoenzymatic approach.[53] The focus of the synthesis is to make the pharmacophores of anticoagulant heparin.[54]

Another chemoenzymatic synthesis used the bacteria *E. coli* K5.[55] This gave an unsulphated precursor of the heparin, heparosan. The chemoenzymatic method came with a lot of limitations, like substrate specificities of the enzyme. The heparin-based oligosaccharide has also been synthesized using the polymer-supported method.[45,56] However, the glycosylation yield enhancement remained a challenge, and hence solution phase is still favoured.

The chemical synthesis of heparin is not like the total synthesis of heparin. Most of the researchers have synthesized the active portion of the chain like pentasaccharides or hexasaccharides as heparin-based oligosaccharide and further evaluated for the corresponding biological activity. This is an excellent approach to get the potential pharmaceutical active molecule with comparatively lesser effort.

MEDICINAL APPLICATIONS OF HEPARIN

Heparin for the first time used as postoperative anticoagulation in 1935, and since then, their applications have increased many folds.[57] The utility of heparin, directly or indirectly, in other therapy like vein thrombosis, heart-lung oxygenation, kidney dialysis, and coating of medical devices like stents make this molecule vital for medicinal applications.

Anticoagulation effect of heparin

A blood clot known as a thrombus is consists of fibrin and blood cells, which can form anywhere in the cardiovascular system.[58,59] The clot can be in the microcirculation, the heart, arteries, and veins. Many diseases are associated with thrombosis including, cardiac-related disease.[60]

Heparin is used for the treatment or prevention of thrombosis, re-thrombosis after thrombolysis, and other thrombosis-related diseases.[10] The blood-clotting cascade is inhibited by a serine proteinase inhibitor, anti-thrombin III (AT).[61] AT is a plasma protein and named by Abildgaard in 1968 as anti-thrombin III, which later referred to as anti-thrombin (AT)[62] after the study of Brinkhous et al. that heparin requires a plasma cofactor for its activity.[63] A pentasaccharide sequence (Figure 4) present in heparin is responsible for the anticoagulation effect.[64,65] The sulfo- and carboxyl groups possessing the negative charge bind tightly with AT.[66,67] AT itself is a weak protease inhibitor, but in combination with heparin pentasaccharide, increase its activity by about 300 times due to favourable conformational change.[68] The formation of insoluble fibrin clots from soluble fibrinogen is facilitated by thrombin. This step is blocked due to the formation of a complex AT-heparin-thrombin, and hence the formation of an insoluble fibrin clot is avoided.[69,70]

As per the mechanism of action, a number of coagulation factors such as XIIa, XIa, IXa, Xa, and IIa (thrombin) are inactivated by the heparin-AT complex.[6,71-73] The unfractionated heparin (UFH) possesses low bioavailability (30%) in comparison to LMWHs (90%) and a longer half-life (17–21 h) when compared to UFH. The LMWHs are given for thromboprophylaxis in fixed doses.[73,74] Due to low bioavailability, the doses need to be high, like 30,000 U/day, when administered via subcutaneous injection for therapeutic anticoagulation.[73]

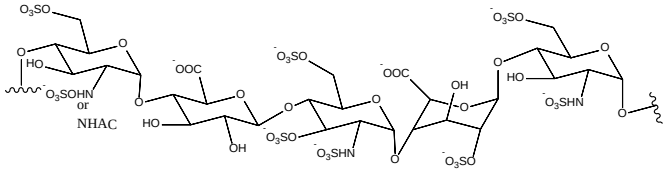


Figure 4. Heparin pentasaccharide that bind to AT

In an ongoing pandemic, COVID-19 also, heparin has been widely explored for its utility, which includes both its anticoagulant effect and antiviral effect.[10] The study showed that anticoagulant therapy with heparin has a positive effect and lowered mortality of COVID-19 patients. The COVID-19 patients with higher D-dimer levels or met the criteria of sepsis-induced coagulopathy (SIC) better impacted with heparin therapy.[75] Patients with 4-times higher D-dimer concentrations than normal were recommended for LMWH.[76] Some studies also confirmed the effectiveness of heparin in COVID patients with venous thromboembolism prophylaxis.[77,78]

Anticancer effect of heparin

The detailed studies of unfractionated and different molecular weight heparin showed improved effects on various types of cancer.[11,79-81] The cancerous tumors and thrombosis showed some link between them; this led to the use of antithrombotic molecules towards the treatment of cancer. The mechanism of action is still a matter of investigation; a noncoagulation pathway is a possible suggestion.[82] The study with heparin on animal models showed a reduction in metastasis of carcinoma cells.[83-85] This may be due to the non-deposition of fibrin around tumur cells.[83] The inhibition of biological processes that supports cell growth is an important approach to target cancer.[86] In some studies, heparin possessing non-anticoagulant property also inhibited metastasis.[84,85]

Several studies confirmed that heparin has effects on angiogenesis and also affects the formation and progression steps of the tumor.[87] They regulate the immune system due to their ability to bind with different proteins. Their role in cell adhesion, a barrier to leukocyte migration, and inflammation have been studied.[11,88] Based on various studies, several heparin-like molecules or heparin mimetics such as pixatimod, muparfostat, necuparanib, and roneparstat are currently in clinical trial for use in cancer treatment.[11] Heparin was used to study its impact on the treatment of angiogenic tumours in mice.[89] The result showed inhibition in downstream signalling and decreased angiogenesis. The heparin, by competition, interrupts the heparin sulphate proteoglycans interaction with fibroblast growth factors (FGF).[89]

The heparin sulphate is present in the extracellular matrix plays several characters towards cell division and proliferation, such as a positive or negative modulator.[90] This has been observed that proliferation of many cells is inhibited by heparin, and this behaviour is utilized for the treatment of cancer.[91] Due to the positive results obtained after many studies for cancer

treatment by heparin or its derivatives, more than 50 clinical trials are at different stages. However, the overall beneficial nature of these drugs is still a matter of investigation.

Antiviral effect of heparin

The GAGs are present on the cell surfaces, which allow them to serve as a non-specific receptor for virus binding.[92] Various studies have been done with heparin against a crowd of distinct viruses.[10,93-95] The human immunodeficiency virus (HIV) showed a reduction with heparin towards cytopathogenicity to MT-4 cells by preventing the adhesion of HIV to MT-4 cells.[96] The heparin-sulphate made the dengue virus ineffective by inhibiting the virus-cell attachment.[97] This has been observed that the heparin blocks the virus adhesion by competitive inhibition.[98] A study by Lin et al. showed the inhibitory action of heparin towards replication of Japanese encephalitis viruses in BHK-21 cells and dengue-2 in hepatoma.[99] The study was done with heparin at various dosages like 0.1, 1, 10, and 100 mg/ml). The result showed that the dengue-2 virus invasion can be stopped by heparin in different liver cells. The influenza virus strain H5N1 infection was also prevented by heparin and its derivatives.[100] The N-sulphation of HS showed its potential towards the infectivity of the Chikungunya virus.[101]

The heparin also showed its effectiveness towards SARS-CoV-2.[10] The angiotensin-converting enzyme 2 (ACE-2) is the major receptor, whereas heparan sulphate proteoglycan (HSPG) is the co-receptor for the virus, SARS-CoV-2, which this virus utilizes to infect the cells. Heparin inhibits the binding of SARS-CoV-2 to the cell surface by competing with co-receptor HSPG, and hence, stops the virus entry to the cell.[10,102] Conzelmann et al. studied SARS-CoV-2 infection inhibition using heparin as a repurposed drug.[103] The study was based on three important therapeutic actions of heparin, anti-inflammatory, anticoagulant, and antiviral. The results showed that the viral replication was completely inhibited with 500–1,000 μg/ml heparin and get suppressed to 60% at 125–250 μg/ml.[103]

Antibacterial activity of heparin

For the first time, Stoker reported the antibacterial effect of heparin against *Staphylococcus aureus*.[104] One year later, another group also reported the antibacterial effect of heparin against *S. aureus* and *Erwinia stewartia*.[105] Hanno et al. in 1978 studied the effect of heparin as an antibacterial agent on rabbit bladder.[106] The mucoprotein-deficient bladder was applied with heparin directly gave a very good result towards bacterial infection. Rosett and Hodges studied the effect of heparin on eight species of microorganisms in brain heart infusion broth.[107] It was observed in all the studies that the microorganism growth was inhibited by heparin. The gram-positive bacteria got inhibited more than the gram-negative bacteria. The unfractionated heparin showed inhibition in a dose-dependent manner to three of seven *S. pneumoniae* isolates and one of five *H. influenzae* isolates.[108] The heparin coating also showed antibacterial properties. The heparin-coated stents prevented the formation of bacterial biofilm and microbe attachment.[109] The biodegradable ureteral stents coated with heparin were also found to inhibit the early bacteriuria development.[110] The above examples confirm the utility of heparin as antibacterial agents. Systematic research is required to bring this as a broad-spectrum antibacterial agent.

Anti-inflammatory activity of heparin

Several studies have been done on heparin as an anti-inflammatory agent.[111-113] Different studies have discussed the mechanism of action; still, exact benefits and safety are yet to be established about heparin being an anti-inflammatory agent. Some of the early studies have suggested the positive use of intravenous heparin in asthma, a chronic inflammatory disorder of the airways.[114,115] The inhalation of heparin also gave positive effect towards bronchial hyperreactivity.[116] The heparin inhalation may be showing the suppressive type of activity on mast cell degranulation which is due to the absence of bronchodilation in airways.[116] In the burn's patients also, heparin was able to reduce the inflammation and also promoted tissue repair.[117]

Heparin and its derivatives showed anti-inflammatory property and benefits to the patients during cataract surgery, cardiopulmonary bypass, and asthma; however, other inflammatory diseases gave mixed and inconsistent results.[113] The adverse effect of heparin as anti-inflammatory agent has not been reported in maximum studies. Still, more studies are required to make this molecule as a versatile and robust anti-inflammatory agent.

HEPARIN-BASED NANOCARRIERS FOR DRUG DELIVERY

Heparin is biocompatible, and hence, heparin-based nanocarriers showed a lot of potential in drug delivery.[118,119] Different types of nanocarriers have been studied with heparin like self-assemblies, coated nanoparticles, nanogels, and polyelectrolyte complex nanoparticles.[119] The size and shape of heparin-based nanoparticles can easily be controlled by optimizing the heparin amount. The physical or chemical interaction using the cross-linked method allows the formation of heparin-based nanogels possessing 3-dimentional (3-D) porous network, which helps in the prevention of environmental degradation of drugs along with drug reservoirs. Chang et al. developed a heparin/berberine conjugate to deliver berberine at the infected site that reduced cytotoxic effects in infected cells.[120] This process reduced the unwanted side-effects caused due to antibiotic treatment for *Helicobacter pylori* bacteria. This was an excellent example of a nanoparticle berberine carrier with a heparin shell.[120] Choi et al. studied the controlled release of growth factors using heparin nanosponge.[121] The growth factors are useful to many diseases therapy and tissue engineering, but due to low bioavailability, non-specific biodistribution, and fast degradation, their efficacy is a limitation. These limitations are reduced with the use of heparin-based nanocarriers.

Heparin-based nanocarriers have been studied for cancer management either as a backbone, coating material, physical encapsulation, and conjugates.[119] An elevated antitumor effect was observed with doxorubicin encapsulation by a conjugate made from heparin and deoxycholic acid that was planned for SCC (squamous cell carcinoma).[122] This was amphiphilic

conjugate nanoparticles and gave a better result than without encapsulation. Zhang et al. studied conjugation and physical loading for two anticancer drugs, D (doxorubicin) and A (all trans retinoic acid) where D was loaded, and A was conjugated with LMWH forming D-loaded LMWH–A system.[123] The study reported that the anticancer effect with the prepared system was higher as compared to the free drugs in solution. Yang et al. studied the applicability of drug sorafenib in gastric cancers using an immobilized chitosan/heparin pluronic-coated system as a nano delivery system.[124,125] This was observed that the prepared nano delivery system enhanced the inhibition of cancer cells.

There have been numerous studies done to use heparin as a nano system for drug delivery, but still, no system has reached to the stage of successful clinical trials. Heparin being the hydrophilic molecule possessing a lot of modifiable groups, can provide an excellent potential in the drug delivery systems. Drug delivery is an essential aspect to provide treatment of various ailments, and a lot of research on varied topics are going on.[126,127]

SUMMARY AND WAY FORWARD

Heparin is a highly sulphated and most negatively charged natural biopolymer belonging to the glycosaminoglycan (GAG) family. This is about a 100-year-old anticoagulant drug, and still, the scientific community unable to find its substitute with similar efficacy. Various bottlenecks towards its isolation from animal sources kept the scientific community at their toes to search for similar alternatives. From publications point of view, many researchers have published its isolation from different sources, chemoenzymatic synthesis, chemical synthesis, biotechnological approach, but still, their commercial applicability is far away. In recent times, the focus is more towards the synthesis of LMWH, ULMWH, and bioengineered heparins.

Heparin is equally important for non-anticoagulant diseases and nanocarriers for drug delivery systems also. The effect of heparin and its oligosaccharides showed a better therapeutic effect towards cell proliferation, inflammation, microbial effect, thrombogenesis, and related diseases. More in-vitro, in-vivo, and clinical studies are required to understand the efficacy and treatment effect of heparin-based oligosaccharides, which are easy to synthesize in comparison to the total synthesis of heparin.

ACKNOWLEDGEMENTS

The authors are grateful to Delhi Technological University (DTU) for providing necessary facilities to carry out this work. The author SJ is also thankful for financial support from DTU.

REFERENCES AND NOTES

1. S.N. Baytas, R.J. Linhardt. Advances in the preparation and synthesis of heparin and related products. *Drug Discov. Today* **2020**, 25, 2095-2109.
2. S.L. Taylor, J. Hogwood, W. Guo, E.A. Yates, J.E. Turnbull. By-products of heparin production provide a diverse source of heparin-like and heparan sulfate glycosaminoglycans. *Sci. Rep.* **2019**, 9, 2679.
3. J.Y. van der Meer, E. Kellenbach, L.J. van den Bos. From farm to pharma: an overview of industrial heparin manufacturing methods. *Molecules* **2017**, 22, 1025–1037.
4. E. Vilanova, B.C. Vairo, S-NMCG Oliveira, B.F. Glauser, N.V. Capillé, G.R.C. Santos, A.M.F. Tovar, M.S. Pereira, P.A.S. Mourão. Heparins sourced from bovine and porcine mucosa gain exclusive monographs in the Brazilian pharmacopeia. *Front. Med.* **2019**, 6:16.
5. W.D. Comper. Heparin (and Related Polysaccharides): Structural and Functional Properties. Gordon and Breach Science Publishers, **1981**.
6. E.I. Oduah, R.J. Linhardt, S.T. Sharfstein. Heparin: Past, present, and future, *Pharmaceuticals* **2016**, 9, 38 (12 pages).
7. C.S. Ong, J.A. Marcum, K.J. Zehr, D.E. Cameron, A Century of Heparin. *Ann. Thorac. Surg.* **2019**, 108, 955-958
8. D. Wardrop, D. Keeling. The story of the discovery of heparin and warfarin. *Br. J. Haematol*. **2008**, 141, 757–763.
9. K.M. Brinkhous, H.P. Smith, E.D. Warner, W.H. Seegers. The inhibition of blood clotting: an unidentified substance which acts in conjunction with heparin to prevent the conversion of prothrombin into thrombin. *Am. J. Physiol.* **1939**, 125, 683–687.
10. C. Shi, T. Wua, J.-P. Li, M.A. Sullivan, C. Wang, H. Wang, B. Deng, Y. Zhang. Comprehensive landscape of heparin therapy for COVID-19. *Carbohydr. Polym.* **2021**, 254, 117232.
11. J. Atallah, H.H. Khachfe, J. Berro, H.I. Assi. The use of heparin and heparin-like molecules in cancer treatment: a review. *Cancer Treat. Res. Commun.* **2020**, 24, 100192.
12. B.L. Farrugia, M.S. Lord, J. Melrose, J.M. Whitelock. The role of heparan sulfate in inflammation, and the development of biomimetics as anti-inflammatory strategies. *J. Histochem. Cytochem.* **2018**, 66, 321–336.
13. J.M. Whitelock, R.V. Iozzo. Heparan sulfate: a complex polymer charged with biological activity. *Chem. Rev.* **2005**, 105, 2745–2764.
14. L. Fu, M. Suflita, R.J. Linhardt. Bioengineered heparins and heparan sulfates. *Adv. Drug Deliv. Rev.* **2016**, 97, 237–249.
15. T. Wang, L. Liu, J. Voglmeir. Chemoenzymatic synthesis of ultralow and low-molecular weight heparins. *Biochim. Biophys. Acta* **2020**, 1868, 140301.
16. J. Kreuger, L. Kjellen. Heparan sulfate biosynthesis: regulation and variability. *J. Histochem. Cytochem.* **2012**, 60, 898–907.
17. D.Y. Lee, S.Y. Lee, H.J. Kang, Y. Park, S.J. Hur. Development of effective heparin extraction method from pig by-products and analysis of their bioavailability. *J. Anim. Sci. Technol.* **2020**, 62, 933-947.
18. D. Keire, B. Mulloy, C. Chase, A. Al-Hakim, D. Cairatti, E. Gray, J. Hogwood, T. Morris, P.A.S. Mourão, M. da Luz Carvalho Soares, et al. Diversifying the global heparin supply chain: Reintroduction of Bovine Heparin in the United States? *Pharm. Technol.* **2015**, 28, 36.
19. M. Petitou, C.A.A. van Boeckel. A synthetic antithrombin III binding pentasaccharide is now a drug! What comes next? *Angew. Chem. Int. Ed.* **2004**, 43, 3118-3133.
20. M. Petitou, J.C. Jacquinet, J. Choay, J.C. Lormeau, M. Nassr. U.S. Patent 4,818,816, **1989**.
21. https://www.cnbctv18.com/healthcare/govt-targets-heparin-for-indigenous-manufacturing -6818461.htm (retrieved on 7th March **2021**)
22. https://thehealthmaster.com/2020/09/04/India-considers-allowing-import-of-api-for-mfg-crude-heparin/ (retrieved on 7th March **2021**)
23. https://economictimes.indiatimes.com/industry/healthcare/biotech/pharmaceuticals/india-considers-allowing-import-of-raw-material-for-manufacturing-crude-heparin/articleshow/ 77903713.cms?utm_source=contentofinterest&utm_ medium=text&utm_campaign=cppst (retrieved on 7th March **2021**).
24. European Pharmacopeia. Monograph for Heparin Sodium Heparinum Natricum, 9th ed.; 01/2017:0333; Council of Europe: Strasbourg, France, **2017**.
25. Guidance for Industry-Heparin for Drug and Medical Device Use: Monitoring Crude Heparin for Quality. Available online: https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulato

ryInformation/Guidances/UCM291390.pdf (accessed on 5th May **2021**).
26. EudraLex. The Rules Governing Medicinal Products in the European Union, Volume 4: EU Guidelines for Good Manufacturing Practice for Medicinal Products for Human Van Veterinary Use, Annex 2 Manufacture of Biological Active Substances and Medicinal Products for Human Use. Ref. Ares (2012) 118531-28/06/2012. Available online: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-4/vol4-an2__2012-06_en.pdf (accessed on 5th May **2021**).
27. C. Casalone, J. Hope. Atypical and classic bovine spongiform encephalopathy. *Handbook of Clinical Neurology* **2018**, 153, 121–134.
28. L. Fu, G. Li, B. Yang, A. Onishi, L. Li, P. Sun, F. Zhang, R.J. Linhardt. Structural characterization of pharmaceutical heparins prepared from different animal tissues. *J. Pharm. Sci.* **2013**, 102, 1447–1457.
29. G. Torri, A.Naggi. Heparin centenary – an ever-young life-saving drug, Int. J. Cardiol. **2016**, 212, S1-S4.
30. P. Bianchini, L. Liverani, G. Mascellani, B. Parma. Heterogeneity of unfractionated heparins studied in connection with species, source, and production processes. *Semin. Thromb. Hemost.* **1997**, 23, 3–10.
31. A. Kouta, W. Jeske, D. Hoppensteadt, O. Iqbal, Y. Yao, J. Fareed. Comparative pharmacological profiles of various Bovine, Ovine, and Porcine Heparins. *Clin. Appl. Thromb. Hemost.* **2019**, 25, 1-9.
32. L.B. Jaques, E.T. Waters, A.F. Charles. A comparison of the heparins of various mammalian species. *J. Biol. Chem.* **1942**, 144, 229-235.
33. R. Einarsson, L.O. Andersson. Binding of heparin to human antithrombin III as studied by measurements of tryptophan fluorescence. *Biochim. Biophys. Acta* **1977**, 490, 104–111.
34. R.D. Rosenberg. Heparin action. *Circulation* **1974**, 49, 603–605.
35. R.J. Linhardt, S.A. Ampofo, J. Fareed, D. Hoppensteadt, J.B. Mulliken, J. Folkmanll. Isolation and characterization of human heparin. *Biochem.* **1992**, 31, 12441-12445.
36. D.D. Metcalfe, R.A. Lewis, J.E. Silbert, R.D. Rosenberg, S.I. Wasserman, K.F. Austen. Isolation and characterization of heparin from human lung. *J. Clin. Invest.* **1979**, 64, 1537-1543.
37. M. Warda, E.M. Gouda, T. Toida, L. Chi, R.J. Linhardt. Isolation and characterization of raw heparin from dromedary intestine: evaluation of a new source of pharmaceutical heparin. *Comp. Biochem. Physiol. C* **2003**, 136, 357–365.
38. M. Warda, W. Mao, T. Toida, R.J. Linhardt. Turkey intestine as a commercial source of heparin? Comparative structural studies of intestinal avian and mammalian glycosaminoglycans. *Comp. Biochem. Physiol. B* **2003**, 134, 189–197.
39. W.P. Jeske, M.K. McDonald, D.A. Hoppensteadt, E.C. Bau, A. Mendes, C.P. Dietrich, J.M. Walenga, E. Coyne. Isolation and characterization of heparin from Tuna Skins. *Clin. Appl. Thromb. Hemost.* **2007**, 13, 137-145.
40. R. Flengsrud, M.L. Larsen, O.R. Ødegaard. Purification, characterization and *in vivo* studies of salmon heparin. *Thromb. Res.* **2010**, 126, e409–e417.
41. C.P. Dietrich, J.F. Paiva, R.A. Castro, S.F. Chavante, W. Jeske, J. Fareed, P.A. Gorin, A. Mendes, H.B. Nader. Structural features and anticoagulant activities of a novel natural low molecular weight heparin from the shrimp *Penaeus brasiliensis*. *Biochim. Biophys. Acta* **1999**, 1428, 273–283.
42. M. Cesaretti, E. Luppi, F. Maccari, N. Volpi. Isolation and characterization of a heparin with high anticoagulant activity from the clam *Tapes phylippinarum*: Evidence for the presence of a high content of antithrombin III binding site. *Glycobiol.* **2004**, 14, 1275–1284.
43. P.H. Seeberger, D.B. Werz. Synthesis and medical applications of oligosaccharides. *Nature*. **2007**, 446, 1046–1051.
44. R. Lucas, J. Angulo, P.M. Nieto, M. Martin-Lomas. Synthesis and structural study of two new heparin-like hexasaccharides. *Org. Biomol. Chem.* **2003**, 1, 2253–2266.
45. R. Ojeda, O. Terenti, J.L. de Paz, M. Martin-Lomas. Synthesis of heparin-like oligosaccharides on polymer supports. *Glycoconj. J.* **2004**, 21, 179–195.
46. J.C. Lee, X.A. Lu, S.S. Kulkarni, Y.S. Wen, S.C. Hung. Synthesis of heparin oligosaccharides. *J. Am. Chem. Soc.* **2004**, 126, 476–477.
47. S. Dey, H.-J. Lo, C.-H. Wong. Programmable one-pot synthesis of heparin pentasaccharide Fondaparinux. *Org. Lett.* **2020**, 22, 4638–4642.
48. P.K. Bundhun, M. Shaik, J. Yuan. Choosing between Enoxaparin and Fondaparinux for the management of patients with acute coronary syndrome: A systematic review and meta-analysis. *BMC Cardiovasc. Disord.* **2017**, 17, 116.
49. M. Mende, C. Bednarek, M. Wawryszyn, P. Sauter, M.B. Biskup, U. Schepers, S. Bräse. Chemical synthesis of glycosaminoglycans. *Chem. Rev.* **2016**, 116, 8193–8255.
50. S. Masuko, R.J. Linhardt. Chemoenzymatic synthesis of the next generation of Ultralow MW Heparin therapeutics. *Future Med. Chem.* **2012**, 4, 289–296.
51. B.L. Farrugia, M.S. Lord, J. Melrose, J.M. Whitelock. Can We Produce Heparin/Heparan Sulfate Biomimetics Using "Mother-Nature" as the Gold Standard? *Molecules*, **2015**, 20, 4254-4276.
52. J. Liu, L.C. Pedersen. Anticoagulant heparan sulfate: structural specificity and biosynthesis. *Appl. Microbiol. Biotechnol.* **2007**, 74, 263-272.
53. Y. Xu, S. Masuko, M. Takieddin, H. Xu, R. Liu, J. Jing, S.A. Mousa, R.J. Linhardt, J. Liu; Chemoenzymatic synthesis of homogeneous ultralow molecular weight heparins. *Science* **2011**, 334, 498-501.
54. D. Loganathan, H.M. Wang, L.M. Mallis, R J. Linhardt. Structural Variation in the Antithrombin III Binding Site Region and Its Occurrence in Heparin from Different Sources. *Biochem.* **1990**, 29, 4362-4368.
55. U. Lindahl, J. Li, M. Kusche-Gullberg, M. Salmivirta, S. Alaranta, T. et al. Generation of "Neoheparin" from E. Coli K5 Capsular Polysaccharide. *J. Med. Chem.* **2005**, 48, 349–352.
56. S.B. Dulaney, X. Huang. Strategies in synthesis of heparin/heparan sulfate Oligosaccharides: 2000–present. *Adv. Carbohydr. Chem. Biochem.* **2012**, 67, 95–136.
57. R.J. Linhardt. Heparin: structure and activity. *J. Med. Chem.* **2003**, 46, 2551–2554.
58. L. Badimon, J.J. Badimon, V. Fuster. Pathogenesis of thrombosis. In: V. Fuster, M. Verstraete, eds. Thrombosis in Cardiovascular Disorders. Philadelphia, Pa: WB Saunders Co. **1992**, 17–39.
59. L.A. Harker, K.G. Mann. Thrombosis and fibrinolysis. In: Fuster V, M. Verstraete. eds. Thrombosis in Cardiovascular Disorders. Philadelphia, Pa: WB Saunders Co. **1992**, 1–16.
60. A. de Laforcade. Diseases associated with thrombosis. *Top. Companion. Anim. Med*. **2012**, 27, 59-64.
61. R.D. Rosenberg, P.S. Damus. The purification and mechanism of action of human antithrombin-heparin cofactor. *J. Biol. Chem.* **1973**, 248, 6490–6505.
62. U. Abildgaard. Highly purified antithrombin 3 with heparin cofactor activity prepared by disc electrophoresis. *Scand. J. Clin. Lab. Invest.* **1968**, 21, 89 –91.
63. K.M. Brinkhous, H.P. Smith, E.D. Warner, W.H. Seegers. The inhibition of blood clotting: an unidentified substance which acts in conjunction with heparin to prevent the conversion of prothrombin into thrombin. *Am. J. Physiol*. **1939**, 125, 683– 687.
64. L. Thunberg, G. Backstrom, U. Lindahl. Further characterization of the antithrombin binding sequence in heparin. *Carbohydr. Res.* **1092**, 100, 393–410.
65. J. Choay, M. Petitou, J.C. Lormeau, P. Sinay, B. Casu, G. Gatti. Structure–activity relationship in heparin: a synthetic pentasaccharide with high affinity for antithrombin III and eliciting high antifactory Xa activity. *Biochem. Biophys. Res. Commun.* **1983**, 116, 492–499.
66. U. Lindahl, G. Backstrom, L. Thunberg, I.G. Leder, Evidence for a 3-O-sulfated d-glucosamine residue in the anti-thrombin binding sequence of heparin. *Proc. Natl. Acad. Sci. USA* **1980**, 77, 6551–6555.

67. L. Jin, J.P. Abrahams, R. Skinner, M. Petitou, R.N. Pike, R.W. Carrell. The anticoagulant activation of antithrombin by heparin. *Proc. Natl Acad. Sci. USA* **1997**, 94, 14683–14688.
68. S.T. Olson, I. Bjork, R. Sheffer, P.A. Craig, J.D. Shore, J. Choay. Role of the anti-thrombin binding pentasaccharide in heparin acceleration of antithrombin–proteinase reactions. Resolution of the antithrombin conformational change contribution to heparin rate enhancement. *J. Biol. Chem.* **1992**, 267, 12528–12538.
69. C.A.A. VanBoeckel, M. Petitou. The unique antithrombin III binding domain of heparin: a lead to new synthetic antithrombotics. *Angew. Chemie. Int. Ed.* **1993**, 32, 1671–1690.
70. Olson ST, Halvorson HR, Bjork I. Quantitative characterization of the thrombin-heparin interaction. Discrimination between specific and nonspecific binding models. *J. Biol. Chem.* **1991**, 266, 6342–6352.
71. J. Hirsh, S.S. Anand, J.L. Halperin, V. Fuster. Guide to Anticoagulant Therapy: Heparin A Statement for Healthcare Professionals from the American Heart Association. *Circulation* **2001**, 103, 2994-3018.
72. D.B. Cines. Heparin: Do we understand its antlthrombotic actlons? *Chest* **1986**, 89, 420-426.
73. M. Alquwaizani, L. Buckley, C. Adams, J. Fanikos. Anticoagulants: A review of the pharmacology, dosing, and complications. *Curr. Emerg. Hosp. Med. Rep.* **2013**, 1, 83–97.
74. T.W. Barrowcliffe. Low molecular weight heparins. *Br. J. Haematol.* **1995**, 90, 1–7.
75. N. Tang, H. Bai, X. Chen, J. Gong, D. Li, Z. Sun. Anticoagulant treatment is associated with decreased mortality in severe coronavirus disease 2019 patients with coagulopathy. *J. Thromb. Haemost.* **2020**, 18, 1094–1099.
76. L. Lin, L. Lu, W. Cao, T. Li. Hypothesis for potential pathogenesis of SARS CoV-2 infection-a review of immune changes in patients with viral pneumonia. *Emerg. Microbes & Infect.* **2020**, 9, 727–732.
77. Z. Zhang, X. Liu, J. Chen, H. Su, Q. Luo, J. Ye, et al. Heparin sulphate D-glucosaminyl 3-O-sulfotransferase 3B1 plays a role in HBV replication. *Virology*. **2010**, 406, 280–285.
78. B. Bikdeli, M.V. Madhavan, D. Jimenez, T. Chuich, I. Dreyfus, E. Driggin, et al. COVID-19 and thrombotic or thromboembolic disease: implications for prevention, antithrombotic therapy, and follow-up: JACC state-of-the-art review. *J. Am. Coll. Cardiol.* **2020**, 75, 2950–2973.
79. R. Alyahya, T. Sudha, M. Racz, S.C. Stain, S.A. Mousa, Anti-metastasis efficacy and safety of non-anticoagulant heparin derivative versus low molecular weight heparin in surgical pancreatic cancer models. *Int. J. Oncol.* **2015**, 46, 1225–1231.
80. R.J.K. Hettiarachchi, S.M. Smorenburg, J. Ginsberg, M. Levine, M.H. Prins, H.R. Buller. Do heparins do more than just treat thrombosis? The influence of heparins on cancer spread. *Thromb. Haemost.* **1999**, 82, 947–952.
81. N.M. Kuderer, A.A. Khorana, G.H. Lyman, C.W. Francis. Ameta-analysis and systematic review of the efficacy and safety of anticoagulants as cancer treatment: impact on survival and bleeding complications. *Cancer* **2007**, 110, 1149–1161.
82. N. Zhang, W. Lou, F. Ji, L. Qiu, B.K. Tsang, W. Di. Low molecular weight heparin and cancer survival: clinical trials and experimental mechanisms. *J. Cancer Res. Clin. Oncol.* **2016**, 142, 1807–1816.
83. D.F. Alonso, G.E. Bertolesi, E.F. Farias, A.M. Eijan, E.B.D. Joffe, L.L. De Cidre. Antimetastatic effects associated with anticoagulant properties of heparin and chemically modified heparin species in a mouse mammary tumor model. *Oncol. Rep.* **1996**, 3, 219–222.
84. S.A. Mousa, R. Linhardt, J.L. Francis, A. Amirkhosravi. Anti-metastatic effect of a non-anticoagulant low-molecular weight heparin versus the standard low-molecular-weight heparin, enoxaparin. *Thromb. Haemost.* **2006**, 96, 816–821.
85. T. Sciumbata, P. Caretto, P. Pirovano, P. Pozzi, P. Cremonesi, G. Galimberti, F. Leoni, F. Marcucci. Treatment with modified heparins inhibits experimental metastasis formation and leads, in some animals, to long-term survival. *Invasion Metastasis.* **1996**, 16, 132–143.
86. N. Nishida, H. Yano, T. Nishida, T. Kamura, M. Kojiro. Angiogenesis in cancer. *Vasc. Health Risk Manag*. **2006**, 2, 213–219.
87. X. Shen, J. Fang, X. Lv, Z. Pei, Y. Wang, S. Jiang, K. Ding, Heparin impairs angiogenesis through inhibition of microRNA-10b. *J. Biol. Chem.* **2011**, 286, 26616–26627.
88. M. Tamim, M. Demircin, M. Guvener, O. Peker, M. Yilmaz. Heparin-coated circuits reduce complement activation and inflammatory response to cardiopulmonary bypass. *Panminerva Med.* **1999**, 41, 193–198.
89. N. Quarto, F. Amalric. Heparan sulphate proteoglycans as transducers of FGF-2 signalling. *J. Cell Sci*. **1994,** 107, 3201–3212.
90. M.A. Porcionatto, H.B. Nader, C.P. Dietrich, Heparan sulfate and cell division. Braz. *J. Med. Biol. Res.* **1999**, 32, 539–544.
91. Y. Uzun, E. Akdogan, F. Ozdemir, E. Ovali. The effects of heparin on DLD-1 colon cancer cell line. *Bratisl. Lek. Listy.* **2009**, 110, 3–6.
92. F. Vigant, N.C. Santos, B. Lee. Broad-spectrum antivirals against viral fusion. *Nat. Rev. Microbiol.* **2015**, 13, 426–437.
93. M. Lima, T. Rudd, E. Yates, New applications of heparin and other glycosaminoglycans. *Molecules*. **2017**, 22, 749 (11 pages).
94. A.L. Howell, T.H. Taylor, J.D. Miller, D.S. Groveman, E.H. Eccles, L.R. Zacharski. Inhibition of HIV-1 infectivity by low molecular weight heparin. *Int. J. Clin. Lab. Res.* **1996**, 26, 124–131.
95. J.A. Hippensteel, W.B. LaRiviere, J.F. Colbert, C.J. Langouët-Astrié, E.P. Schmidt. Heparin as a therapy for COVID-19: current evidence and future possibilities. *Am. J. Physiol. Lung Cell Mol. Physiol.* **2020**, 319, L211–L217.
96. M. Ito, M. Baba, A. Sato, R. Pauwels, E. De Clercq, S. Shigeta. Inhibitory effect of dextran sulfate and heparin on the replication of human immunodeficiency virus (HIV) in vitro. *Antiviral Res.* **1987**, 7, 361–367.
97. N. Modhiran, N.S. Gandhi, N. Wimmer, S. Cheung, K. Stacey, P.R. Young, et al. Dual targeting of dengue virus virions and NS1 protein with the heparan sulfate mimic PG545. *Antiviral Res.* **2019**, 168, 121–127.
98. M. Sasaki, P.D. Anindita, N. Ito, M. Sugiyama, M. Carr, H. Fukuhara, T. Ose, K. Maenaka, A. Takada, W.W. Hall, Y. Orba, H. Sawa. sThe role of heparan sulfate proteoglycans as an attachment factor for rabies virus entry and infection. *J. Infect. Dis.* **2018**, 217, 1740–1749.
99. Y.L. Lin, H.Y. Lei, Y.S. Lin, T.M. Yeh, S.H. Chen, H.S. Liu. Heparin inhibits dengue-2 virus infection of five human liver cell lines. *Antiviral Res.* **2002**, 56, 93–96.
100. M.A. Skidmore, A. Kajaste-Rudnitski, N.M. Wells, S.E. Guimond, T.R. Rudd, E.A. Yates, E. Vicenzi. Inhibition of influenza H5N1 invasion by modified heparin derivatives. *Med. Chem. Commun.* **2015,** 6, 640–646.
101. A. Tanaka, U. Tumkosit, S. Nakamura, D. Motooka, N. Kishishita, T. Priengprom, et al. Genome-wide screening uncovers the significance of N-sulfation of heparan sulfate as a host cell factor for Chikungunya virus infection. *J. Virology* **2017**, 91, 1–22.
102. S.E. Guimond, C.J. Mycroft-West, N.S. Gandhi, J.A. Tree, K.R. Buttigieg, N. Coombes, et al. Pixatimod (PG545), a clinical-stage heparan sulfate mimetic, is a potent inhibitor of the SARS-CoV-2 virus. *bioRxiv*. **2020**, 2020.2006.2024.169334.
103. C. Conzelmann, J.A. Müller, L. Perkhofer, K.M.J. Sparrer, A.N. Zelikin, J. Münch, A. Kleger. Inhaled and systemic heparin as a repurposed direct antiviral drug for prevention and treatment of COVID-19. *Clin. Med.* **2020**, 20, e218–21
104. S. B. Stoker. Bacteriostatic action of heparin. *J. Physiol (London)*. **1949**, 1l0, 26P.
105. J.R. Warren, F. Graham. The effect of heparin on the growth of bacteria and yeasts. *J. Bacteriol.* **1950**, 60, 171-174.
106. P.M. Hanno, R. Fritz, A.J. Wein, S.G. Mulholland. Heparin as antibacterial agent in Rabbit bladder. *Urology* 1978, XII, 411-415.
107. W. Rosett, G.R. Hodges. Antimicrobial Activity of Heparin. *J. Clin. Microbiol.* **1980**, 11, 30-34.

108. C. Zappala, S. Chandan, N. George, J. Faoagali, R.J. Boots. The antimicrobial effect of heparin on common respiratory pathogens. *Crit. Care Resusc.* **2007**, 9, 157-160.
109. P. Tenke, C.R. Riedl, G.L. Jones, G.J. Williams, D. Stickler, E. Nagy. Bacterial biofilm formation on urologic devices and heparin coating as preventive strategy. *Int. J. Antimicrob. Agents* **2004**, 23, S67–S74.
110. J.E. de la Cruz Conty, A.B. Alba, J.L.S. Miguelanez, J.A.G. Llopis, T.F. Aparicio, F.M.S. Margallo, F.S. Galvez. Heparin as an antibacterial coating on biodegradable ureteral stents: experimental study of Braidstent-H. *BJS*, **2021**, 108(Supplement 3) iii1–iii20.
111. R.J. Ludwig. Therapeutic use of heparin beyond anticoagulation. *Curr. Drug Discov. Tech.* **2009**, 6, 281–289.
112. E. Young. Theanti-inflammatory effects of heparin and related compounds. *Thromb. Res.* **2008**, 122, 743–752.
113. S. Mousavi, M. Moradi, T. Khorshidahmad, M. Motamedi, Anti-inflammatory effects of heparin and its derivatives: A systematic review. *Adv. Pharmacol. Sci*. **2015**, 2015, Article ID 507151, 14 pages.
114. J.P. Boyle, R.H. Smart, J.K. Shirey. Heparin in the treatment of chronic obstructive bronchopulmonary disease. *Am. J. Card.* **1964**, 14, 25–28.
115. N.L. Fine, C. Shim, M.H. Williams Jr. Objective evaluation of heparin in the treatment of asthma. *Am. Rev. Respirat. Dis.* **1968**, 98, 886–887.
116. T. Ahmed, J. Garrigo, I. Danta. Preventing bronchoconstriction in exercise-induced asthma with inhaled heparin. *New Engl. J. Med.* **1993**, 329, 90–95.
117. J.M.F. Chacon, M.L.M. De Andrea, L. Blanes, L.M. Ferreira. Effects of topical application of 10,000 IU heparin on patients with perineal dermatitis and second-degree burns treated in a public pediatric hospital. *J. Tissue Viab*. **2010**, 19, 150–158.
118. G. Baier, S. Winzen, C. Messerschmidt, D.F.M. Fichter, S. Gehring, V. Mailander, K. Landfester. Heparin-based nanocapsules as potential drug delivery systems. *Macromol. Biosci.* **2015**, 15, 765–776.
119. X. Yang, H. Du, J. Liu, G. Zhai. Advanced Nanocarriers Based on Heparin and Its Derivatives for Cancer Management. *Biomacromol.* **2015**, 16, 423–436.
120. C.H. Chang, W.Y. Huang, C.H. Lai, Y.M. Hsu, Y.H. Yao, T.Y. Chen, J.Y. Wu, S.F. Peng, Y.H. Lin. Development of novel nanoparticles shelled with heparin for berberine delivery to treat Helicobacter pylori. *Acta Biomaterialia*. **2011**, 7, 593–603.
121. W.I. Choi, A. Sahu, C. Vilos, N. Kamaly, S.M. Jo, J.H. Lee, G. Tae. Bioinspired heparin nanosponge prepared by photo-crosslinking for controlled release of growth factors. *Sci Rep*. **2017**, 7, 14351.
122. K. Park, G. Y. Lee, Y.-S. Kim et al., Heparin-deoxycholic acid chemical conjugate as an anticancer drug carrier and its antitumor activity. *J. Cont. Release.* **2006,** 114, 300–306.
123. T. Zhang, H. Xiong, F. Z. Dahmani, L. Sun, Y. Li, L. Yao, J. Zhou, J. Yao. Combination chemotherapy of doxorubicin, all-trans retinoic acid and low molecular weight heparin based on self-assembled multifunctional polymeric nanoparticles. *Nanotech.* **2015**, 26, 145101.
124 Y.C. Yang, J. Cai, J. Yin, J. Zhang, K. L. Wang, Z.-T. Zhang. Heparin-functionalized Pluronic nanoparticles to enhance the antitumor efficacy of sorafenib in gastric cancers. *Carbohydr. Polym.* **2016**, 136, 782–790.
125 R. Patel, B. Gajra, R. Parikh, G. Patel. Preparation, characterization and in-vitro drug release evaluation of Ganciclovir loaded Chitosan nanoparticles. *J. Mater. NanoSci*., **2016**, 3(2), 41–48.
126 P. Mittal, S. Singh, A. Singh, I.K. Singh. Current advances in drug delivery systems for treatment of triple negative breast cancer (TNBC). *Chem. Biol. Lett.* **2020**, 7, 1-12.
127 P. Kumari, M. Gulia, S. Gupta, D. Singh, S. Kumar. An economic and efficient synthesis of acid-labile glycerol based β-thiopropionate esters for potential application in drug delivery. *Chem. Biol. Lett.* **2021**, 8, 45-49.

AUTHORS BIOGRAPHIES



Saikrushna Jena was born in Balasore, Odisha, India. She received her B.Sc. in Chemistry, M.Sc. in Polymer Science from Central Institute of Plastic Engineering and Technology, India in 2017. At present, she is pursuing her M.Tech. in Polymer Technology under supervision of Dr. Ram Singh at Delhi Technological University, Delhi, India. Her research interest lies in biopolymer and waste management.



Dr. Ram Singh was born in Chandauli, Sitamarhi, Bihar, India. He received his B.Sc. (Chem, Hons), M.Sc. (Org. Chem.) and Ph.D. (Chemistry) from University of Delhi, India. During post-doctoral period, he worked at CEMDE, University of Delhi and Fast Track Young Scientists (DST) at Department of Chemistry, University of Delhi. At present, he is working at Delhi Technological University since 2010. He is working in the field of bioorganic chemistry, natural products, and waste management. His research work has been funded by DST, CSIR and DRDO, India. He is on the Editorial Advisory Board of various journals of repute and Life Member of various societies.

The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/0972-9968.htm

Life dissatisfaction among students: exploring the role of intrapersonal conflict, insufficient efforts and academic stress

Anju, Amandeep, B.K. Punia and Vandana Punia
Guru Jambheshwar University of Science and Technology, Hisar, India, and
Naval Garg
Delhi Technological University, Delhi, India

Received 27 September 2020
Revised 6 December 2020
Accepted 21 December 2020

Abstract

Purpose – The study focuses on academic stress among the students due to insufficient efforts and intrapersonal conflicts amid the COVID-19 situation. The purpose of this study is to measure the mediating role of academic stress between life dissatisfaction and adequate steps and intrapersonal conflicts.

Design/methodology/approach – The researcher had surveyed 729 students from the two states, including Haryana and Punjab and 716 responses were used for analysis. These respondents were approached in April–May 2020 to explore the impact of COVID-19. The questionnaire was prepared with the help of Google form, including 29 questions. Confirmatory factor analysis and structural equation modeling (SEM) were used to validate the research model and for testing the hypothesis.

Findings – The findings indicate that intra-personal conflict and insufficient efforts have a significant positive relation with academic stress. Also, intrapersonal conflict and insufficient efforts have a significant positive relation with life dissatisfaction. It has been reported that academic stress mediates the relationship between life dissatisfaction and intra-personal conflict. The results also reveal that academic stress is positively and significantly associated with life dissatisfaction.

Originality/value – This is one of the few studies that explored the mediating role of academic stress between life dissatisfaction and insufficient efforts and intrapersonal conflicts.

Keywords Academic stress, Life dissatisfaction, Insufficient efforts, Intrapersonal conflicts

Paper type Research paper

Introduction

In today's world, students have to compete at every step-in academic resulting in a lot of stress. This growing academic pressure has resulted in "Academic stress." Researchers in recent literature have studied academic stress in university students (Gross, 2014, 2015a, 2015b; Freire *et al.*, 2018; Phang *et al.*, 2015). Chao (2012) shares that stress in university students can be a cumulative emotion, resulting due to academic uncertainty of the future, doubting self, difficulties in forming interpersonal- social relationships and many other

© Anju, Amandeep, B.K. Punia, Vandana Punia and Naval Garg. Published in *Rajagiri Management Journal*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence maybe seen at http://creativecommons.org/licences/by/4.0/legalcode

Rajagiri Management Journal
Emerald Publishing Limited
0972-9968
DOI 10.1108/RAMJ-09-2020-0058

reasons. Gupta and Khan (1987) defined it as a "mental distress concerning some anticipated frustration associated with academic failure or even awareness of the possibility of such failure."

Though stress is "normative" in the developmental period, there are times when this stress has been found to persist later in life also. Situations, where extreme stress is experienced by the student, can be fatal. Bhujade (2017) shared that at any given time, 10–20% of students are suffering from some form of mental health problem in Indian colleges. These mental health problems may be an outcome of increased stress and may appear in the form of anxiety and depression (Kumaraswamy, 2013). Suffering from mental health issues while pursuing academics may be dysfunctional for the student. Also, these stress levels influence one's satisfaction with university and life in general (Jenkins *et al.*, 2013).

Academic stress is a generalized phenomenon, it can be experienced at school and college (university) level. Students may experience it independent of their academic performance or capability. Academically weaker students may be worried about their performance in class, while others are worried about maintaining or bringing up their grades. Conditions in the family such as divorce, intrapersonal conflicts and marital depression can affect the functioning of the student (Forehand *et al.*, 1991). An unsupportive environment at home/school causes a feeling of burden among students, which leads to suicidal ideation (Arun and Chavan, 2009). Also, students' personal expectations to excel and that of their parents and teachers contribute to the building up of academic stress (Ang and Huan, 2006; Deb *et al.*, 2014).

It is interesting to note that there are differences in the experience of stress in female and male students, especially in the Indian context. Verma and Verma (2020) have shared that women experience stress due to lack of support and men are burdened due to the expectations of the parents. They feel that as their parents are supporting them financially and emotionally (friends, family, relatives, etc.), they in return may have some expectations such as good academic scores, better career and achievements. This belief in a student's mind builds pressure that as he or she is getting all the facilities and support from the surrounding, he/she has to perform better to fulfill these expectations. This increased pressure and deteriorating mental health of students on campus is a concern as that affects the psychological health of students and that of future adults. Recent work reveals that academic stress should be kept under control for students to have a rewarding experience at the university (Freire *et al.*, 2018).

The current paper focuses on two variables that can be the reason for academic stress in student life. These two variables are personally related to an individual and vary according to his or her personality type. These cognitive factors influence the level of stress experienced, leading to the satisfaction experienced by the student toward life. It is important to note that there are prior personal characteristics that influence the experience of stress in students (de la Fuente, 2015; Craig *et al.*, 2020).

The first variable in the study is insufficient effort. Here the term efforts mean the strength of a student that he or she puts into performing better in their academic life. Insufficient efforts can become the reason for poor performance. It directly affects the stress level in a student's life. Not only a weak student but a brilliant student also gets deviated from his path and is not able to focus on his goal, thus affecting the level of effort they put into their task. With increased competition in academics, insufficient efforts can make your performance lag. All these things are the main reasons for academic stress in a student's life.

Another variable we have used here is intrapersonal conflicts. Bazerman *et al.* (1998) identify intrapersonal conflict as a discrepancy between what they want to do and what they should do. The positive and negative thoughts and feelings which one would hold toward an attitude or object create intrapersonal conflicts. The contradictions which the students

experience internally could be due to motivation to do something or conflicts of self-esteem or many other factors (Gerasimova and Gerasymova, 2019). These differences which internal self-experiences with the external environment contribute to conflicts are experienced by them. For intrapersonal conflicts to arise, there is a need for a personal and social condition to be present and academics can be seen as one of the conditions. The cognitive dissonances experienced by the individual can contribute to the development of intrapersonal conflicts. In the current study, we are discussing conflict within oneself. When students have some uncertain feelings about themselves and their capabilities, they find themselves in a dilemma. They start to doubt their capabilities, which becomes a hurdle in their performance. As the dilemma increases in their life, it directly impacts their academic performance and increases the stress experienced. As stress directly impacts an individual's life, in the same way, academic stress negatively affects the student's life. In this current study, the researchers also focus on the negative impact of academic stress on students' satisfaction level in life. The current paper is focusing on the relationship between stressors of academic stress and life's dissatisfaction.

Literature review

Academic stress among students

Academic stress has been defined as a negative and undesirable response of students toward academic environmental stressors (Lee and Larson, 2000). It also includes students' appraisal, analysis and reactions to studies, institutions and career-related stressors. Owing to the gravity of issues and challenges posed by academic stress, it has been termed as a "career stopper" (Kadapatti and Vijayalaxmi, 2012). Subramani and Kadhiravan (2017) proposed a linkage between academic stress and mental health. The authors clarified that students often get cramped and caged in academic structure and they hardly find any route to deal with the academic stressors. Nandamuri and Gowthami (2011) argued that academic stress is one of the main reasons for the exponential rise in mental health issues among Indian youth. Sharma *et al.* (2016) opined that academic stress might lead to students' mental and physical challenges. Deb *et al.* (2014) and Beiter *et al.* (2015) reported that academic stress might be observed in the form of depression, anxiety, intolerance, anger, irritability, loss of concentration and focus and withdrawal symptoms in students.

In some cases, students resort to extreme deviant behaviors such as suicide, run-away from home, drug addiction break institutional rules and procedures (Garg and Saxena, 2020). Huang *et al.* (2020) study how uncivil behavior and academic stress influence the psychological health of the student. Busari (2012) suggested that a stressed mind and body make it difficult for a student to concentrate. It fills the student with negative emotions and subsequently, an unknown fear of failure surrounds the student. Several researchers, including Saxena *et al.* (2020), reported that excessive stressful adolescent students are more prone to indulge in alcoholism and drug abuse. Few other undesirable repercussions of stressed academic life are physical and mental inactivity, dullness, eating and sleeping habit disorder and unprotected and failed sexual life.

Studies have shown that a mild level of stress is conducive for the student and that lowering the level of stress does not ensure better performance. A student may consider a situation with no stress as non-challenging and the performance may reduce significantly (Uchil, 2017). Thus, practitioners and academicians opined that a certain threshold level of stress might encourage students to perform better. Caution is that prolonged stress or inadequately managed stress negatively influences students' prospects and academic institutions. Mishra and Mckean (2000) argued that, apart from stressors, lack of synergy and synchronization of stressors and students' approach to cope with these stressors create tension, anxiety and depression among students.

Researchers and academicians agree that prevailing challenging conditions emerging out of pandemic COVID-19 have aggravated students' academic stress (Bhat *et al.*, 2020). Students' anxiety level has shot-up due to uncertainty over classes and examinations, physical distancing, longer screen time, poor connectivity in rural areas, lack of laptops or desktops, etc. Pandemic has resulted in reduced motivation and morale of students toward studies, increased pressures to learn independently, desertion of daily routines, reduced approachability of faculty members and potentially higher dropout rates (Grubic *et al.*, 2020). Sheroun *et al.* (2020) reported that academic stress took an inverted "U" shaped graph in the lockdown period. Academic stress was lower in the first 13 days of lockdown; then, it raised sharply for the next 15 days. There on, it decreased after one month, but this level was quite higher than the first 13 days (Sheroun *et al.*, 2020). These studies highlight the importance of understanding academic stress during the COVID-19 situation.

Interpersonal conflict, insufficient efforts and academic stress

Previous researchers have discussed several reasons for academic stress among students. These reasons include homesickness, fear of failure, the pressure of studies, competition, financial hardship, etc. Zeidner (1992) elaborated that there are physical conditions, economic factors, a psychological weakness, etc., which contribute to academic stress. Detailing further on, Kohn and Frazer (1986) observed 35 academic stressors for students. These stressors were the uncertainty of final grade, examination, boring classes, parents and self-expectations, etc.

Amid these multiplicities of academic stressors, interpersonal conflict and insufficient efforts are being explored for their probable role in students' academic stress. The present study explores the effect of these two stressors on students for the following reasons. Intrapersonal conflicts contribute to stress, independent of other stresses present in the environment (Lauterbach, 1991). Mikolaj and Boggs (1991) studied how intrapersonal conflicts contribute to stress in adult students. Researchers studying academic stress argued that roles that require a higher level of interpersonal- interactions are more stressful. "Interpersonal problems. . .ranks as a primary source of unhappiness in people's lives. A construct that measures the quality of interpersonal relationships at work is interpersonal conflict" (Jaramillo *et al.*, 2011). Then, a student must develop interpersonal skills to interact with teachers, colleagues and other stakeholders of a college. It is noteworthy to understand that individual personality differences may give rise to interpersonal conflict among students. Interpersonal conflict is defined as negative interactions with co-workers, supervisors or clients, which leads to heated arguments, unfriendly or violent behavior. Among students, interpersonal conflict may manifest as unhealthy competition, back-bitting, criticism, ostracism, quarrel and grouping. It is interesting to note that most interpersonal conflicts lead to petty arguments and gossip, not physical and violent attacks (Schat, Frone and Kelloway, 2006). Keenan and Newton (1985) suggested that interpersonal conflicts elicit anger, frustration and stress in people who encounter them. It has also been argued that interpersonal conflicts are more detrimental to mental health than non-interpersonal conflicts (Taniguchi and Kato, 2018). Students have reported that the institution community's social interactions are among the most challenging parts of their school careers (Pyhältö *et al.*, 2010). The supportive relationships with school friends and teachers play a pivotal role in students' well-being and successful adaptation to educational context (Longobardi *et al.*, 2016). Several other researchers such as De Dreu *et al.* (2004), Jamadin *et al.* (2015), Karatepe and Baddar (2006), Applebaum *et al.* (2010), Wang *et al.* (2018), Fuller *et al.* (2003) and Miller (1989) supported the view that interpersonal conflict

leads to stress. Stress negatively influences the academic achievement of university students.

Insufficient efforts represent the students' lack of willpower and low-level efforts to resolve the problems, leading to stress (Fuller *et al.*, 2003; Miller, 1989). A student may find it difficult to devote equal efforts to studies, relationship management and physical and mental health well-being. Misdirected efforts, imbalanced efforts and insufficient efforts hinder goal achievement, eliciting anger and stress among students (Jamadin *et al.*, 2015). The following hypotheses were proposed to explore the proposed relationships between academic stress, interpersonal conflict and insufficient efforts.

Life dissatisfaction, interpersonal conflict, insufficient efforts and academic stress

Life satisfaction is one of the most vibrant constructive and positive adjustment indicators influenced by students' stress during their academic life (Çivitci, 2015). Life satisfaction has been considered one of the three components of subjective well-being (Çivitci, 2015). The other two tenants of subjective well-being, positive and negative, consider pleasant and unpleasant emotions only. However, life satisfaction talks about the cognitive evaluation of individuals' lives (Diener and Suh, 1997). Thus, life satisfaction is a much more enduring, stable, reliable and important factor of subjective well-being because it allows long-term evaluations (Eid and Diener, 2004). Apart from being an indicator of subjective well-being, life satisfaction also indicates the objective quality of life such as health, financial status, freedom and entertainment (Diener and Suh, 1997).

Life satisfaction has been associated with college success and has been found to contribute to limiting mental health symptoms and low stress (Jenkins *et al.*, 2013). However, increased college stress has been associated with lower life satisfaction (Brougham *et al.*, 2009). Ng *et al.* (2020), in their research, established the relation of intrapersonal factors and life satisfaction among youth. They share that higher intrapersonal conflicts lead to lower life satisfaction, especially for young adults.

Burger and Samuel (2017), in their longitudinal study, demonstrated that internally stronger students are able to better deal with stress during difficult transitions in life. These cognitive factors, if they have strength, mitigate the negative effect of stress, thereby leading to greater life satisfaction.

Ben-Zur (2009), in his study, shared that well-being is associated with social support and coping. Parental support and teacher support and counseling play an important role (Virtanen *et al.*, 2019) in providing that strength to the students in dealing with stress. In this study, the negative connotation of life satisfaction, i.e. life dissatisfaction, was explored. The authors proposed that interpersonal conflict and insufficient efforts lead to life dissatisfaction among students. These relationships are also proposed to be mediated by academic stress. Accordingly, the following model (Figure 1) and hypotheses were proposed. Figure 2 represent results of the hypothesized model..

H1. There is a significant relationship between insufficient efforts and academic stress.

H2. There is a significant relationship between insufficient efforts and life dissatisfaction.

H3. There is a significant relationship between intrapersonal conflicts and academic stress.

H4. There is a significant relationship between intrapersonal conflicts and life dissatisfaction.

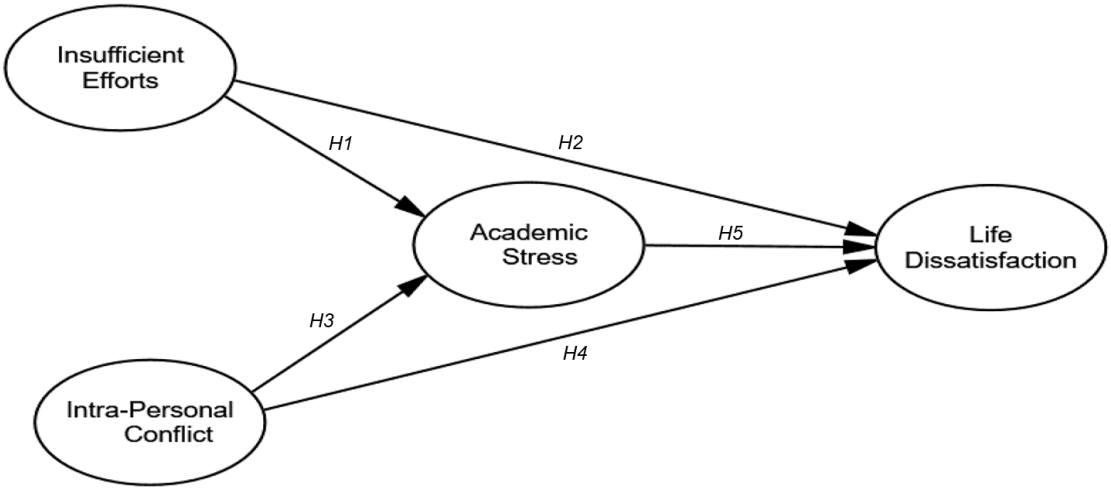


Figure 1.
Hypothesized model

Source: Author's Own

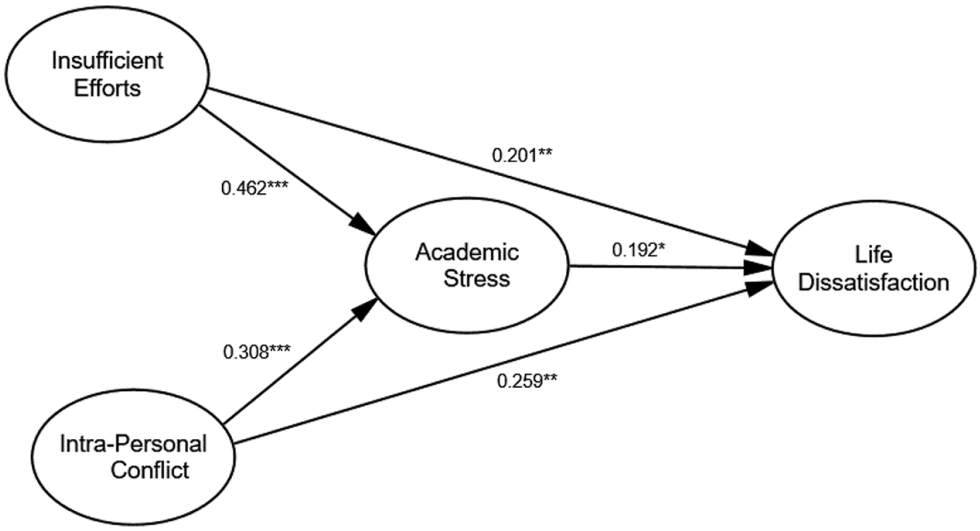


Figure 2.
Model results

Source: Primary data

H5. There is a significant relationship between academic stress and life dissatisfaction.

H6. The positive relationship between intrapersonal conflicts and life dissatisfaction is mediated by academic stress.

H7. The positive relationship between insufficient efforts and life dissatisfaction is mediated by academic stress.

Methodology and hypothesis development

Method

Data collection and the sample. The survey was conducted on college students from two states- Haryana and Rajasthan. The questionnaire was prepared on Google forms, which included 35 questions. In all a sample of 729 students was collected, of which 14 responses

were excluded where 10 were found to be not engaged and 3 had duplicity. Hence, a sample of 716 was used for analysis. In the final sample, 472 (65.9%) were female and 244 (34.1%) were male. There were, 509 (71.1%) respondents pursuing graduation, 174 (24.3%) post-graduation and 33 (4.6%) were enrolled in Ph.D. Out of 716 respondents, 82.1% were living in their residence, 8.2% stayed in college/university hostel, 5.3% in relative house, 3.6% were living in a rented room and 0.7% in PG. Of the 716 respondents, 54.7% belonged to rural areas and 45.3% were from urban areas.

Measures

To measure academic stress 7-items scale was used from a 21-item inventory of academic stress (Lakaev, 2009) and 2 items were dropped due to low factor loading. To measure the insufficient efforts and intrapersonal conflicts, an 11-item scale was used from the Level of attribution and change (LAC) scale (Norcross *et al.*, 1985). All the items of Satisfaction with life scale (SWLS) were modified for measuring life dissatisfaction (Diener *et al.*, 1985) and 3 items were used for this scale and 2 items were dropped due to low factor loadings. We measured all of the items on a five-point Likert scale ranging from "strongly disagree" (1) to "strongly agree" (5).

Normality, reliability, validity and common method variance estimates

We observed a fairly normal distribution for our indicators of latent factors in terms of skewness. However, we observed mild kurtosis for the indicators of our dependent variable life satisfaction (LS), mediator academic stress (ACS) and independent variables interpersonal conflicts and Insufficient efforts (IPC and IE). The kurtosis values ranged from benign to 1.898. While this does violate strict rules of normality, it is within more relaxed rules suggested by Sposito *et al.* (1983), who recommended 3.3 as the upper threshold for normality. We ran cook's distance analysis to determine if any (multivariate) influential outlier existed. In no case did we observe a cook's distance greater than 1. Most cases were far less than 0.025. Cronbach's α values of all the variables were more than the 0.70 threshold value considered acceptable by Cortina (1993) revealing satisfactory reliability of all the constructs. The Composite Reliability (CR) of each scale exceeded the value of 0.70 thresholds for acceptable reliability (Bagozzi and Yi, 1988; Nunnally, 1994) and that the AVE value exceeded the recommended threshold of 0.5 (Bagozzi and Yi, 1988; Fornell and Larcker, 1981). Finally, as shown in Table 1, the AVE's square root exceeded all the correlations (Fornell and Larcker, 1981), demonstrating sufficient discriminant validity between the constructs. Also, Heterotrait-Monotrait ratio (HTMT) criteria were adopted to prove discriminant validity between constructs. All the ratios were less than the $HTMT_{85}$ criteria (Kline, 2011), demonstrating sufficient discriminant validity between the constructs

Variables	Mean	SD	α	1	2	3	4
1. ACS	2.57	0.955	0.792	*0.739*			
2. IE	2.44	1.019	0.863	0.590***	*0.804*		
3. IPC	2.23	0.954	0.872	0.571***	0.704***	*0.781*	
4. LDS	2.09	0.724	0.727	0.385***	0.439***	0.451***	*0.803*
5. CR				0.856	0.901	0.903	0.844

Notes: SD, standard deviation; ACS, academic stress; IPC, inter-personal conflict; IE, insufficient efforts; LDS, life dissatisfaction; CR, composite reliability; α, Cronbach's alpha; diagonal italic values represent ✓AVE; ***$p < 0.001$

Table 1.
Correlation matrix

Direct effects	Standardized coefficient	*t*-value
Intra-personal conflict → academic stress	0.308	3.853***
Insufficient efforts → academic stress	0.462	5.464***
Intra-personal conflict → life dissatisfaction	0.259	2.597**
Insufficient efforts → life dissatisfaction	0.201	2.784**
Academic stress → life dissatisfaction	0.192	2.048*
Indirect effects	***p*-value**	**Standardized coefficient**
Intra-personal conflict → academic stress → life dissatisfaction	0.014	0.051 (0.011–0.111)
Insufficient efforts → academic stress → life dissatisfaction	0.017	0.072 (0.013–0.154)

Note: ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. Numbers in brackets represent the 95% bias-corrected and accelerated confidence intervals derived from bootstrapping with 2,000 samples
Source: Primary data

Table 2.
Mediation results

as shown in Table 3. CFA was used to examine the reliability and validity of study measures. The results revealed that the measurement model fitted the data well for study: χ^2 (141) = 376.754, χ^2/df = 2.67, GFI = 0.947, CFI = 0.962, TLI = 0.933, RMSEA = 0.048 (Hair *et al.*, 1998; Hu and Bentler, 1999). Previous research indicates a good fit of the model and the data if the RMSEA value is less than 0.05 (NE and Cudeck, 1993). Further, the value of standardized RMR is 0.0363. All the factor loads were significant and had a value of over 0.5 (Stevens, 2012).

As we collected the data of dependent and independent variables from the same respondent through self-reporting, there could be common method variance (CMV). Harman's single factor test explained total variance explained of 41.036% by a single factor, so there is no common method bias in our data set. We also conducted a confirmatory factor analysis to compare the four-factor model with a single-factor model in which all items were loaded on a single factor (Podsakoff *et al.*, 2003). According to the fit indices (χ^2 = 1,398.042, df = 152, $p = 0.000$; GFI = 0.801; TLI = 0.772, CFI = 0.797; RMSEA = 0.107), the single-factor model provided an unacceptable fit to the data, implying that CMV was not a serious issue.

Results

Table 1 shows the correlations between all the variables and correlations were all in the expected direction. The findings demonstrate that academic stress is positively correlated with insufficient efforts ($r = 0.716$, $p < 0.001$), intra-personal conflicts ($r = 0.691$, $p < 0.001$) and life dissatisfaction ($r = 0.513$, $p < 0.001$). Life dissatisfaction is positively correlated with intra-personal conflict ($r = 0.551$, $p < 0.001$) and insufficient efforts ($r = 0.549$, $p < 0.001$).

Constructs	1	2	3	4
1. Academic stress				
2. Insufficient efforts	0.699			
3. Intra-personal conflicts	0.663	0.800		
4. Life dissatisfaction	0.482	0.538	0.559	

Source: Primary data

Table 3.
Assessment of discriminant validity using the HTMT

To analyze the validity of our proposed hypotheses, we used SEM procedures with maximum likelihood estimation. The structural model analysis indicated a good fit, as judged from the goodness-of-fit indices $\chi 2 = 375.372$; $\chi 2/df = 2.662$; $p < 0.001$; CFI = 0.962; GFI = 0.949; AGFI = 0.931; RMSEA = 0.048; RMR = 0.07; SRMR = 0.0369.

The findings, as depicted in Table 2, indicate that intra-personal conflict and insufficient efforts are positively related to academic stress ($\beta_{IPC \rightarrow ACS} = 0.308$, $p < 0.001$ and ($\beta_{IE \rightarrow ACS} = 0.462$, $p < 0.001$, respectively). Therefore, *H1* and *H3* are supported. In addition, intrapersonal conflict and insufficient efforts also have significant positive relation with life dissatisfaction ($\beta_{IPC \rightarrow LDS} = 0.259$, $p < 0.01$, $\beta_{IE \rightarrow LDS} = 0.201$, $p < 0.05$, respectively). Hence, *H2* and *H4* are supported. In *H3*, we hypothesized that academic stress positively affects life dissatisfaction. The results reveal that academic stress is positively and significantly associated with life dissatisfaction ($\beta_{\mathrm{ACS} \rightarrow \mathrm{LDS}} = 0.192$, $p < 0.01$). Hence, *H5* is supported. Moreover, the bootstrapping analysis suggests that academic stress mediates the relationships between intra-personal conflicts ($\beta_{\mathrm{IPC} \rightarrow \mathrm{ACS} \rightarrow \mathrm{LDS}} = 0.051$, $p < 0.05$), insufficient efforts ($\beta_{\mathrm{IE} \rightarrow \mathrm{ACS} \rightarrow \mathrm{LDS}} = 0.072$, $p < 0.05$) and life dissatisfaction. Thus, *H6* and *H7* are supported.

Discussion

The current study is conducted to measure the effects of insufficient efforts and intrapersonal conflicts on students' academic stress and life dissatisfaction. The results of the study established academic stress as a mediating variable between the two personal factors and life dissatisfaction. In the period of this pandemic, the academic life of every student is becoming more stressful as compared to the pre-pandemic period. The current paper discusses the two variables that cause academic stress and life dissatisfaction among students. *H1* assessed the relationship between insufficient efforts and academic stress; the result of the study supported the hypothesis with a standardized coefficient of 0.308 at a 0.1% level of significance – the student exhibiting insufficient efforts to experience academic stress. Due to the present pandemic condition, students are experiencing new situations every day. These situations deviate from the normal and many of these are beyond the control of the student. Thus, making it difficult for the student to give their best effort and thereby experiencing greater academic stress (Jamadin *et al.*, 2015). The students may lack the capability or assume that they lack the capability to do the desired task. To further see the effect of insufficient efforts, *H3* proposed a relation between insufficient effort and life dissatisfaction. The result was supported with a standardized coefficient of 0.462 at a 0.1% level of significance. The students who experience insufficient efforts can also experience greater dissatisfaction regarding life. These new challenging times may leave the students demotivated, leading to a lower level of effort toward their academic goals (Miller, 1989). Their doubts about competence lowered willpower and not working hard enough can be accentuated by the ongoing pandemic. It is important to note that students in the past have not been exposed to such difficult challenges. They do not have the mental strength or experience to deal with negativity, thus leading to greater life dissatisfaction.

H2 assessed the relationship between intrapersonal conflicts and academic stress. The result of the study confirmed the positive relationship between the two variables with a standardized coefficient of 0.201 at a 0.1% level of significance. The students experience more stress when their intrapersonal conflicts are on the rise. The internal discordance of the students in these unprecedented times would be on the rise. There would have difficulty in experiencing their true feeling, with uncertainty about self and lack of knowledge on how the self should behave in the present situations. The students do not know how to present their inner self in this condition, which has led them to a state of cognitive dissonance,

thereby increasing their level of stress (Gerasimova and Gerasymova, 2019). Further on, *H4* assessed the influence of intrapersonal conflicts on life dissatisfaction and the result of the study confirmed the relationship with a standardized coefficient of 0.259 at a 0.1% level of significance. The greater the internal conflicts are for the students, the more dissatisfied they are with their life. Contemporary work by supports the findings for youth where intrapersonal conflicts are inhibitive of students to experience satisfaction toward life. These internal conflicts can lead students to experience mental health symptoms (Jenkins *et al.*, 2013).

The findings of the study also support *H5*, which had proposed a relation between academic stress and life satisfaction with a standardized coefficient value of 0.192 at a 5% level of significance. The increased stress which the student experiences will have an impact on their satisfaction toward life. Studies (Bhat, 2020 and Sheroun *et al.*, 2020) have established that this pandemic has contributed to increasing stress level of the students with a change in an academic set up such as uncertainty over classes and examinations, physical distancing, longer screen time, poor connectivity in rural areas and lack of laptops or desktops. Our findings are supporting the studies which have established the association between academic stress and life dissatisfaction (Brougham *et al.*, 2009). The study in their *H6* and *H7* proposed to see the indirect effect of insufficient effect and intrapersonal conflict on life dissatisfaction mediated through academic stress. The role of academic stress was established as a mediating variable.

The findings of the study highlight the critical role of student's personal experiences and how they can influence academic and personal well-being. These internal difficulties can lead to a negative influence on the academic performance of the student and could deter his progress in his career (Kadapatti and Vijayalaxmi, 2012). As a result, have shared the mounting intrapersonal conflicts and insufficient efforts lead to academic stress and life dissatisfaction, which can result in mental health problems or symptoms (Sharma *et al.*, 2016; Deb *et al.*, 2014). Subramani and Kadhiravan (2017) had established how the student finds it difficult to deal with the academic stressors and they are unable to find ways to deal with them. Academic failure and dissatisfaction with life can also lead to other forms of deviant behaviors (Saxena *et al.*,2020).

For the well-being of the youth, reducing these individual effects of the personal factors would be critical to reducing the mounting life satisfaction. The ongoing pandemic surely has contributed to the challenges the students are facing. Students might not be able to share them with anyone, as they are restricted in their home environment. Studies (Kong *et al.*, 2015) have established that developing a support system for the student can be influential in lower life dissatisfaction experienced by the students. The support system which is provided by the school, teacher, classmates and family members is critical for satisfaction at school and that can influence life satisfaction (Danielsen *et al.*, 2009). The support from friends and family acts as a protective shield for the students. Due to limited support in these times, establishing a system to protect the student would be useful for dealing with some of the internal challenges which the students are experiencing.

A psychological support system can help deal with the personal challenges of the student. These challenges at an internal level such as self-efficacy, self-esteem, capability issues, complexes should be dealt with by professional experts. Psychologists and counselors introduced into the system can help the student to deal with these issues. Studies have shared that students have an increased perception of stress when they have dysfunctional coping strategies (Chao, 2012). Establishing these psychological support systems can rectify these dysfunctional coping with correct life

skills for personal development. Also, the university should provide information about career-related issues and future related prospects to help with the uncertainty experienced in the present times. There is a positive relationship between social support systems and reducing stress and increasing life satisfaction (Coffman and Gilligan, 2002). Developing a safe, supportive environment would reduce academic stress and also moderate the capability of students to recover from these difficulties (Wilks, 2008).

Practical implication

Our study has important implications for the academic stress that students experience at the university level. It is crucial to control factors that contribute to and build academic dis-stress for students. The personal and internal challenges which the student experiences should be addressed in a non-threatening and supportive environment. First, the university should look into developing a system that addresses these issues with professional experts (psychologists and counselors) through facilities of counseling and support systems (career counseling and guidance, etc.) at the university level. Emotion-focused coping strategies through these support systems would be useful in dealing with mental health challenges (Brougham *et al.*, 2009). Second, establishing a mentor-mentee program has been found to be useful in the past (Bhujade, 2017); thus, focusing on developing such a buddy relationship can be useful in university settings. Developing such close bonds can be useful to closely monitor the health of the students and look into the personal challenges early on which the students might experience. Thirdly, establishing career and counseling and guidance centers at the university level for the student to find one point source for authentic information for their career-related concerns. In the current times of COVID-19, with increased stress, it is a dire need to address the internal issues of the students to strengthen their mental well-being and increase life satisfaction and success at university.

Limitation and future research

The present study has limitations, which should be kept in consideration. The study was conducted in the environment of coronavirus pandemic. The external environment could influence the results of the study, and hence should be generalized with caution. Second, the data was collected through an online survey, the researcher was not able to contact the respondent directly. The third limitation of the study is that it excludes school students. Further research, including school students, may show some different results. Fourth the study has been conducted in the northern and central regions of the country, which does not represent the whole country due to regional and cultural differences. This can also be limited to the generalization of the study. The study does not look at the differences in academic stress for students in different fields and specialization, which also gives scope for further research.

Based on the finding of the current study, some points are recommended for further research. First, the presence of social support and family support/environment can be an important factor to contribute as a moderator to the variables explored in the present study. Second, exploring academic stress in school students can be useful, especially in the board class students (10th and 12th) as their personal factors can be critical in contributing to stress and life dissatisfaction. Another dimension to look at is when we deconstruct academic stress with respect to the stress of the teaching and learning process and personal factors, which contribute to academic stress especially in the Indian context.

Conclusion

The study confirmed the proposed hypothesis that there is an effect of insufficient efforts and intrapersonal conflict on academic stress. The increased academic stress increases life dissatisfaction among the students at the university level. These internal influencers of intrapersonal conflicts and insufficient conflicts influence the life dissatisfaction of the students. The study highlights the need to focus on the psychological needs of the student as increased discordance can develop into academic stress and influence the general meaning of life in a negative way. These challenging times of COVID may bring greater internal discrepancies; universities and colleges should help in creating a supportive environment with services and systems that can help students tide over these difficult times.

References

Ang, R.P. and Huan, V.S. (2006), "Academic expectations stress inventory: development, factor analysis, reliability, and validity", *Educational and Psychological Measurement*, Vol. 66 No. 3, pp. 522-539.

Applebaum, D., Fowler, S., Fiedler, N., Osinubi, O. and Robson, M. (2010), "The impact of environmental factors on nursing stress, job satisfaction, and turnover intention", *JONA: The Journal of Nursing Administration*, Vol. 40 Nos 7/8, pp. 323-328, doi: 10.1097/NNA.0b013e3181e9393b.

Arun, P. and Chavan, B. (2009), "Stress and suicidal ideas in adolescent students in Chandigarh", *Indian Journal of Medical Sciences*, Vol. 63 No. 7, p. 281, doi: 10.4103/0019-5359.55112.

Bagozzi, R.P. and Yi, Y. (1988), "On the evaluation of structural equation models", *Journal of the Academy of Marketing Science*, Vol. 16 No. 1, pp. 74-94.

Bazerman, M.H., Tenbrunsel, A.E. and Wade-Benzoni, K. (1998), "Negotiating with yourself and losing: making decisions with competing internal preferences", *Academy of Management Review*, Vol. 23 No. 2, pp. 225-241.

Beiter, R., Nash, R., McCrady, M., Rhoades, D., Linscomb, M., Clarahan, M. and Sammut, S. (2015), "The prevalence and correlates of depression, anxiety, and stress in a sample of college students", *Journal of Affective Disorders*, Vol. 173, pp. 90-96.

Ben-Zur, H. (2009), "Coping styles and affect", *International Journal of Stress Management*, Vol. 16 No. 2, p. 87.

Bhat, R., Singh, V.K., Naik, N., Kamath, C.R., Mulimani, P. and Kulkarni, N. (2020), "COVID 2019 outbreak: the disappointment in indian teachers", *Asian Journal of Psychiatry*, Vol. 50, p. 102047.

Bhujade, V.M. (2017), "Depression, anxiety and academic stress among college students: a brief review", *Indian Journal of Health and Wellbeing*, Vol. 8 No. 7.

Brougham, R.R., Zail, C.M., Mendoza, C.M. and Miller, J.R. (2009), "Stress, sex differences, and coping strategies among college students", *Current Psychology*, Vol. 28 No. 2, pp. 85-97.

Burger, K. and Samuel, R. (2017), "The role of perceived stress and self-efficacy in young people's life satisfaction: a longitudinal study", *Journal of Youth and Adolescence*, Vol. 46 No. 1, pp. 78-90.

Busari, A.O. (2012), "Evaluating the relationship between gender, age, depression and academic performance among secondary school students", *International Journal of Interdisciplinary and Multidisciplinary Studies*, Vol. 4 No. 2, pp. 6-12.

Chao, R.C.L. (2012), "Managing perceived stress among college students: the roles of social support and dysfunctional coping", *Journal of College Counseling*, Vol. 15 No. 1, pp. 5-21.

Çivitci, A. (2015), "Perceived stress and life satisfaction in college students: belonging and extracurricular participation as moderators", *Procedia-Social and Behavioral Sciences*, Vol. 205, pp. 271-281.

Coffman, D.L. and Gilligan, T.D. (2002), "Social support, stress, and self-efficacy: effects on students' satisfaction", *Journal of College Student Retention: Research, Theory and Practice*, Vol. 4 No. 1, pp. 53-66.

Cortina, J.M. (1993), "What is coefficient alpha? An examination of theory and applications", *Journal of Applied Psychology*, Vol. 78 No. 1, p. 98.

Craig, K., Hale, D., Grainger, C. and Stewart, M.E. (2020), "Evaluating metacognitive self-reports: systematic reviews of the value of self-report in metacognitive research", *Metacogn. Learn*, pp. 1-59, doi: 10.1007/s11409-020- 09222- y.

Danielsen, A.G., Samdal, O., Hetland, J. and Wold, B. (2009), "School-related social support and students' perceived life satisfaction", *The Journal of Educational Research*, Vol. 102 No. 4, pp. 303-320, doi: 10.3200/JOER.102.4.303-320.

De Dreu, C.K., Van Dierendonck, D. and Dijkstra, M.T. (2004), "Conflict at work and individual well-being", *International Journal of Conflict Management*, Vol. 15 No. 1.

de la Fuente, J., Zapata, L., Martínez-Vicente, J.M., Sander, P. and Putwain, D. (2015), "Personal self-regulation, self-regulated learning and coping strategies, in university context with stress", in Pena-Ayala, A. and Cardenas, L. (Eds), *Metacognition: Fundaments, Applications, and Trends*, Vol. 76, Springer, London, pp. 39-72.

Deb, S., Strodl, E. and Sun, J. (2014), "Academic-related stress among private secondary school students in India", *Asian Education and Development Studies*, Vol. 3 No. 2.

Diener, E. and Suh, E. (1997), "Measuring quality of life: economic, social, and subjective indicators", *Social Indicators Research*, Vol. 40, pp. 189-216.

Diener, E.D., Emmons, R.A., Larsen, R.J. and Griffin, S. (1985), "The satisfaction with life scale", *Journal of Personality Assessment*, Vol. 49 No. 1, pp. 71-75.

Eid, M. and Diener, E. (2004), "Global judgments of subjective well-being: situational variability and long-term stability", *Social Indicators Research*, Vol. 65 No. 3, pp. 245-277.

Forehand, R., Wierson, M., Thomas, A.M., Armistead, L., Kempton, T. and Neighbors, B. (1991), "The role of family stressors and parent relationships on adolescent functioning", *Journal of the American Academy of Child and Adolescent Psychiatry*, Vol. 30 No. 2, pp. 316-322.

Fornell, C. and Larcker, D.F. (1981), "Evaluating structural equation models with unobservable variables and measurement error", *Journal of Marketing Research*, Vol. 18 No. 1, pp. 39-50.

Freire, C., Ferradás, M., Núñez, J., Valle, A. and Vallejo, G. (2018), "Eudaimonic well-being and coping with stress in university students: the mediating/moderating role of self-efficacy", *International Journal of Environmental Research and Public Health*, Vol. 16 No. 1, p. 48.

Fuller, J.A., Stanton, J.M., Fisher, G.G., Spitzmüller, C., Russell, S.S. and Smith, P.C. (2003), "A lengthy look at the daily grind: time series analysis of events, mood, stress, and satisfaction", *Journal of Applied Psychology*, Vol. 88 No. 6, pp. 1019-1033.

Garg, N. and Saxena, A. (2020), "Analyzing the inter-relation between workplace spirituality and constructive deviance", *Asian journal of Business Ethics*, Vol. 9, pp. 121-141.

Gerasimova, N. and Gerasymova, I. (2019), "The influence of intrapersonal conflicts on the course of adaptation of students at the establishment of higher education", *Sustainable Development under the Conditions of European Integration*, Vol. Part I, p. 384.

Gross, J.J. (2014), *Handbook of Emotion Regulation*, 2nd ed, Guilford, New York, NY.

Gross, J.J. (2015a), "Emotion regulation: current status and future prospects", *Psychological Inquiry*, Vol. 26 No. 1, pp. 1-26.

Gross, J.J. (2015b), "The extended process model of emotion regulation: elaborations, applications, and future directions", *Psychological Inquiry*, Vol. 26 No. 1, pp. 130-137.

Grubic, N., Badovinac, S. and Johri, A.M. (2020), "Student mental health in the midst of the COVID-19 pandemic: a call for further research and immediate solutions", *International Journal of Social Psychiatry*, p. 0020764020925108.

Gupta, K. and Khan, B.N. (1987), "Anxiety level as a factor in concept formation", *J. Psychol. Reports*, Vol. 31, pp. 187-192.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E. and Tatham, R.L. (1998), *Multivariate Data Analysis*, Vol. 5, Prentice hall Upper Saddle River, NJ.

Hu, L. and Bentler, P.M. (1999), "Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives", *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 6 No. 1, pp. 1-55.

Huang, N., Qiu, S., Alizadeh, A. and Wu, H. (2020), "How incivility and academic stress influence psychological health among college students: the moderating role of gratitude", *International Journal of Environmental Research and Public Health*, Vol. 17 No. 9, p. 3237.

Jamadin, N., Mohamad, S., Syarkawi, Z. and Noordin, F. (2015), "Work–family conflict and stress: evidence from Malaysia", *Journal of Economics, Business and Management*, Vol. 3 No. 2, pp. 309-312.

Jaramillo, F., Mulki, J.P. and Boles, J.S. (2011), "Workplace stressors, job attitude, and job behaviors: is interpersonal conflict the missing link?", *Journal of Personal Selling and Sales Management*, Vol. 31 No. 3, pp. 339-356.

Jenkins, S.R., Belanger, A., Connally, M.L., Boals, A. and Durón, K.M. (2013), "Firstgeneration undergraduate students' social support, depression, and life satisfaction", *Journal of College Counseling*, Vol. 16 No. 2, pp. 129-142.

Kadapatti, M.G. and Vijayalaxmi, A.H.M. (2012), "Stressors of academic stress – a study on pre-university students", *Indian Journal of Scientific Research*, Vol. 3 No. 1, pp. 171-175.

Karatepe, O.M. and Baddar, L. (2006), "An empirical study of the selected consequences of frontline employees' work–family conflict and family–work conflict", *Tourism Management*, Vol. 27 No. 5, pp. 1017-1028.

Keenan, A. and Newton, T.J. (1985), "Stressful events, stressors and psychological strains in young professional engineers", *Journal of Organizational Behavior*, Vol. 6 No. 2, pp. 151-156.

Kline, R.B. (2011), "Methodology in the social sciences", *Principles and Practice of Structural Equation Modeling*, (3rd ed.), Guilford Press.

Kohn, J.P. and Frazer, G.H. (1986), "An academic stress scale: identification and rated importance of academic stressors", *Psychological Reports*, Vol. 59 No. 2, pp. 415-426.

Kong, F., Ding, K. and Zhao, J. (2015), "The relationships among gratitude, self-esteem, social support and life satisfaction among undergraduate students", *Journal of Happiness Studies*, Vol. 16 No. 2, pp. 477-489.

Kumaraswamy, N. (2013), "Academic stress, anxiety and depression among college students: a brief review", *International Review of Social Sciences and Humanities*, Vol. 5 No. 1, pp. 135-143.

Lakaev, N. (2009), "Validation of an australian academic stress questionnaire", *Australian Journal of Guidance and Counselling*, Vol. 19 No. 1, pp. 56-70.

Lauterbach, W. (1991), "Intrapersonal conflict, life stress, and emotion", *Stress and Anxiety*, Vol. 13, pp. 85-92.

Lee, M. and Larson, R. (2000), "'The korean 'examination hell': long hours of studying, distress, and depression", *Journal of Youth and Adolescence*, Vol. 29 No. 2, pp. 249-271.

Longobardi, C., Prino, L.E., Marengo, D. and Settanni, M. (2016), "Student-teacher relationships as a protective factor for school adjustment during the transition from middle to high school", *Frontiers in Psychology*, Vol. 7, p. 1988.

Mikolaj, E.L. and Boggs, D.L. (1991), "Intrapersonal role conflicts of adult women undergraduate students", *The Journal of Continuing Higher Education*, Vol. 39 No. 2, pp. 13-19.

Miller, B. (1989), "Adult children's perceptions of caregiver stress and satisfaction", *Journal of Applied Gerontology*, Vol. 8 No. 3, pp. 275-293.

Mishra, R. and Mckean, M. (2020), "College students' academic stress and its relation to their anxiety, time management, and leisure satisfaction", *American Journal of Health Studies*, Vol. 16 No. 1, pp. 41-51.

Nandamuri, P. and Gowthami, C. (2011), "Sources of academic stress – a study on management students", *Journal of Management and Science*, Vol. 1 No. 2, pp. 31-42.

Ne, M.W. and Cudeck, R. (1993), "Alternative ways of assessing model fit", *Testing Structural Equation Models*, Vol. 154, p. 136.

Ng, J.C., Lau, V.C. and Chen, S.X. (2020), "Why are dispositional enviers not satisfied with their lives? An investigation of intrapersonal and interpersonal pathways among adolescents and young adults", *Journal of Happiness Studies*, Vol. 21 No. 2, pp. 525-545.

Norcross, J.C., Prochaska, J.O. and Hambrecht, M. (1985), "Levels of attribution and change (Lac) scale: development and measurement", *Cognitive Therapy and Research*, Vol. 9 No. 6, pp. 631-649.

Nunnally, J.C. (1994), *Psychometric Theory 3E*, Tata McGraw-Hill Education.

Phang, C.K., Mukhtar, F., Ibrahim, N., Keng, S.L. and Sidik, S.M. (2015), "Effects of a brief mindfulness-based intervention program for stress management among medical students: the mindful-gym randomized controlled study", *Advances in Health Sciences Education*, Vol. 20 No. 5, pp. 1115-1134.

Podsakoff, P.M., MacKenzie, S.B., Lee, J.Y. and Podsakoff, N.P. (2003), "Common method biases in behavioral research: a critical review of the literature and recommended remedies", *Journal of Applied Psychology*, Vol. 88 No. 5, pp. 879-903.

Pyhältö, K., Soini, T. and Pietarinen, J. (2010), "Pupils' pedagogical well-being in comprehensive school – significant positive and negative school experiences of finnish ninth graders", *European Journal of Psychology of Education*, Vol. 25 No. 2, pp. 207-221.

Saxena, A., Garg, N., Punia, B.K. and Prasad, A. (2020), "Exploring role of indian workplace spirituality in stress management: a study of oil and gas industry", *Journal of Organizational Change Management*, Vol. 33 No. 5.

Schat, A.C., Frone, M.R. and Kelloway, E.K. (2006), "Prevalence of workplace aggression in the US workforce: findings from a national study".

Sharma, B., Kumar, A. and Sarin, J. (2016), "Academic stress, anxiety, remedial measures adopted and its satisfaction among medical students: a systematic review", *International Journal of Health Sciences and Research*, Vol. 6, pp. 368-376.

Sheroun, D., Wankhar, D.D., Devrani, A., Lissamma, P.V., Gita, S. and Chatterjee, K. (2020), "A study to assess the perceived stress and coping strategies among B.Sc. nursing students of selected colleges in Pune during COVID-19 pandemic lockdown", *International Journal of Science and Healthcare Research*, Vol. 5 No. 2, pp. 280-288.

Sposito, V.A., Hand, M.L. and Skarpness, B. (1983), "On the efficiency of using the sample kurtosis in selecting optimal lpestimators", *Communications in Statistics – Simulation and Computation*, Vol. 12 No. 3, pp. 265-272.

Stevens, J.P. (2012), *Applied Multivariate Statistics for the Social Sciences*, Routledge.

Subramani, C. and Kadhiravan, S. (2017), "Academic stress and mental health among high school students", *Indian Journal of Applied Research*, Vol. 7 No. 5, pp. 404-406.

Taniguchi, H. and Kato, T. (2018), "The frequencies and effects of interpersonal stress coping with different types of interpersonal stressors in friendships on mental health and subjective well-being among college students", *The Japanese Journal of Personality*, pp. 27-33.

Uchil, H.B. (2017), "The effect of stress on students' performance", *Stress Management Professional International Journal*, Vol. 5 No. 1, pp. 17-21.

Verma, G. and Verma, H. (2020), "Model for predicting academic stress among students of technical education in India", *International Journal of Psychosocial Rehabilitation*, Vol. 24 No. 04, pp. 2702-2714.

Virtanen, T.E., Vaaland, G.S. and Ertesvåg, S.K. (2019), "Associations between observed patterns of classroom interactions and teacher wellbeing in lower secondary school", *Teaching and Teacher Education*, Vol. 77, pp. 240-252.

Wang, D., Wang, X. and Xia, N. (2018), "How safety-related stress affects workers' safety behavior: the moderating role of psychological capital", *Safety Science*, Vol. 103, pp. 247-259.

Wilks, S.E. (2008), "Resilience amid academic stress: the moderating impact of social support among social work students", *Advances in Social Work*, Vol. 9 No. 2, pp. 106-125.

Zeidner, M. (1992), "Sources of academic stress: the case of first year Jewish and Arab college students in Israel", *Higher Education*, Vol. 24 No. 1, pp. 25-40.

Further reading

Calaguas, G.M. (2012), "Survey of college academic stressors: development of a new measure", *Journal of Human Sciences*, Vol. 9 No. 1, pp. 441-457.

Corresponding author

Naval Garg can be contacted at: naval.garg@dtu.ac.in

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**

Indian Journal of Chemical Technology
Vol. 28, March 2021, pp. 139-149

Lipase catalyzed esterification of Docosahexaenoic acid (DHA) with immobilized *Pseudomonas cepacia* and *Thermomyces lanuginosus*

Aditi Sharma[1,2], Kriti Bhandari[1,2,3], Anjali Jain[1,2], Satyendra P. Chaurasia*[,1,2] & Ajay K. Dalai[2]

[1]Chemical Engineering Department, Malaviya National Institute of Technology, Jaipur 302017.

[2] Department of Chemical & Biological Engineering, 57 Campus Drive, College of Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada.

[3]Department of Biotechnology, Delhi Technological university, Delhi 110 085

E-mail: spchaurasia.chem@mnit.ac.in, chch1962@gmail.com

Received 17 April 2019; accepted 28 January 2021

The selective behaviour of lauryl alcohol to catalyze esterification for the synthesis of tuna free fatty acids (TFFAs) using immobilized *Pseudomonas cepacia* (PCL) and *Thermomyces lanuginosus* lipases (TLL) have been investigated. The characterization of both immobilized PCL and TLL and their support immobead-150 is carried out using particle size analyzer, BET method and FT-IR. The suitable reaction conditions have been determined as lauryl alcohol to TFFAs ratio 4:1 (wt/wt), *p*H 8.0, buffer to TFFAs ratio 1:1 (wt/wt), temperature 50°C and agitation speed 800 rpm. Reusability of both the immobilized lipases for esterification has also been carried out and the activity was stabilized at 60.9% and 47.6% after 5 cycles of repeated use of immobilized PCL and TLL, respectively. Maximum 93.8 wt% Docosahexaenoic acid is recovered in esters with immobilized PCL after 16 h.

Keywords: DHA, Esterification, Lauryl alcohol, Lipases, Tuna free fatty acids

The potential of DHA rich food have been widely accepted around the world in terms of its ability to control and prevent different chronic disorders[1-8]. Due to their wide applications, it is crucial to extract highly pure form of DHA cost effectively[9,10]. As reported by various researchers, the major steps involved in the extraction of DHA are hydrolysis and esterification using selective and non-selective applications of different lipase[11-17]. For the esterification of free fatty acids, DHA can be separated either in fatty acids or in esters, depending upon the type of lipase used. For example, lipases from *Candida rugosa*[18,19], *Rhizopus oryzae*[20,21] and *Rhizopus delemar*[22] have shown less selectivity for the DHA during esterification; hence DHA remains in fatty acid phase, whereas lipases from *Candida antarctica* lipase-B, and *Aspergillus niger* etc. have selectivity for DHA esterification[23,24]. Therefore, the selection of suitable lipase is important in both methods of esterification reaction. Furthermore, type of alcohol can also equally affect the extent of reaction and lipase selectivity. The association of lipase selectivity with type of alcohols used in the esterification reaction has been observed by different groups. Shimada *et al*.[22] have noted that the presence of long chain fatty alcohol enhances esterification of fatty acids. In 1998, Shimada *et al*.[25] have reported that lipase can recognize long chain fatty alcohol as a substrate more easily than other alcohols. Again in 2001, Shimada *et al*.[26] found that *Candida antarctica* lipase-B (CAL-B) can also recognize polyunsaturated fatty acids and short chain alcohols as substrate. Relative selectivity and discriminative power of lipases among the series of fatty acids have been characterized in terms of specific alcohol acceptors used in the studies reported by Chang *et al*.[27]. Later, the conversion of oleic acid was correlated with type of alcohol and after 1 h, 100% conversion of oleic acid was observed with 1-butanol and pentanol.

Therefore in the present work, the selective behavior of immobilized PCL and TLL lipases have been studied for a range of fatty acids with various alcohols to improve the separation of DHA. Simultaneously, study for activation energy and reusability of lipases have also been performed for the esterification of DHA selectively in iso-octane medium.

Experimental Section

Materials and methods

Substrate and lipases

The tuna free fatty acids (TFFAs) extracted from tuna fish oil by hydrolysis with immobilized *Candida antarctica* lipase-B (CAL-B) as reported by

Sharma *et al.*[23]. Hence, TFFAs were utilized as a substrate for the selective esterification reaction. For the selective esterification studies, *Pseudomonas cepacia* lipase (Lot No.: 54327; activity ≥900 U/g), *Thermomyces lanuginosus* lipase (Lot No.: 76546; activity ≥3000 U/g), *Rhizopus oryzae* lipase (Lot No.: 89445; activity ≥300 U/g[a]) immobilized on immobead-150 and *Mucor miehei* (Lot No.: 62350; activity ≥30 U/g[b]) immobilized on ion-exchange were purchased from Sigma-Aldrich, Saskatoon (SK), Canada. Immobead-150 (94129) was also purchased from Sigma-Aldrich, Canada to compare their physiochemical properties with the immobilized lipases used in the present work.

Standards of Tridocosahexaenoin (T-310), Didocosahexaenoin (D-311), Monodocosahexaenoin (M-314) and Trieicosahexaenoin (T-325) were purchased from Nu-Check-Prep., Elysian, MN for the analysis and standardization of HPLC with ELSD detector.

Experimental Procedure

The separation of DHA from other fatty acids has been conducted by the selective esterification of tuna free fatty acids (TFFAs) with immobilized PCL and TLL enzymes. The optimization study for esterification reaction was carried out with immobilized PCL for various alcohols (ethanol, 2-propanol, butanol, hexanol, heptanol and lauryl alcohol), alcohol amount of 0.5-3.5 g, *p*H (4-9), buffer amount of 0-1.75 g, speed of agitation of 200-1000 rpm, and temperature of 20-60°C. The optimization experiments were performed in 20 ml screw-capped glass vials (VWR, Canada make) with 0.25 g of tuna free fatty acids as substrate, 1:1 (wt/wt) iso-octane solvent and 250 U (0.27 g) of immobilized PCL for selective esterification of fatty acids in 6 h. The activation energy for the reaction was calculated for both immobilized PCL and TLL at determined optimized parameters by plotting a curve between ln k vs. 1/T as shown in equation (1):

$$\ln k = -\frac{E}{R}\left[\frac{1}{T}\right] + \ln A \quad \ldots (1)$$

where k is the rate constant, A_o is the pre-exponential function, T is temperature in Kelvin, E is the activation energy measured in (kJ/mol) and R is the gas constant. The reaction was stopped by filtering out the lipase from the reaction mixture after every sample collection which was diluted with 3 mL of iso-octane for the GC injection to analyze the results.

The extraction of DHA rich lauryl ester and the release of DHA rich free fatty acids have been carried out after the selective esterification. Then, reusability study was also conducted up to five repeated uses of immobilized PCL and TLL at optimized reaction conditions. According to equation (2), the activity retention is defined as the %conversion achieved after reuse of immobilized enzyme with respect to % conversion achieved after first round of its use. The activity retention can also be expressed as ratio's of %conversion as given in equation (2):

$$\%\ \text{Activity retention for (n+1) cycle} = \frac{\%\ \text{Conversion after (n+1) time use of immobilized enzyme}}{\%\ \text{Conversion after one time use of immobilized enzyme}} \times 100 \quad \ldots (2)$$

All the results were repeated in 3 sets of experiments for performing the reproducibility analysis. This process was performed with a set of saponification and solvent extraction method, a standard procedure reported by Wanasundara and Shahidi[28] and Senanayake and Shahidi[29] with slight modifications.

Analysis

The activity of lipase has been expressed as μmoles of fatty acid esters formed per mg of immobilized lipase per min. After esterification of fatty acids, the mixture containing fatty acid esters and free fatty acids have been analyzed by HPLC (Agilent 1200 Series model) with electron light scattering detector, and an Agilent based PL gel organic GPC column 100A°, 7.5 × 300 mm 5μm column (p/n PL 1110-6520) protected with guard column and equipped with Chemstation for LC3D was used. The area percentage of ester peak was taken as the percentage of esterification (%E) and the area percentage representing FFAs accounts for the wt% of FFA in the mixture. The experiments were performed in repeated sets and a variation in TFFAs ester formation <±5% was observed. An Agilent make gas chromatography system (model 7890A) has been used, equipped with flame ionization detector (FID, 260°C) and capillary column DB-23 (dimensions: 60 m length, 0.25 mm ID, 0.25 μm film) for analysis of free fatty acids in methyl ester form (FAMEs). The initial composition of TFFAs was analyzed by converting the free fatty acids into methyl ester form with BF_3/methanol

solution under inert N_2 purging atmosphere. Throughout the experimentation, the GC was operated at constant conditions (Carrier: hydrogen gas with flow rate 20 cm/min & 23.148 psi pressure; Oven: 140 to 240°C at 4°C/min. and Injection: 1µL sample, 260°C & split: 20:1). The surface properties of a sample are estimated based upon the amount of nitrogen gas absorbed in relationship with its pressure, at the boiling temperature of liquid nitrogen under normal atmospheric pressure. The specific surface area, pore volume, and average pore diameter for immobilized lipases and their support materials immobead-150 and ion-exchange resin, were measured by N_2-physisorption at 77 K using a Micro metrics ASAP 2000. Approximately 0.2-0.3 g of sample was used for each analysis. The moisture and other adsorbed gases present in the sample were removed before analysis by degassing the sample at 60°C for 1 h under 66.7 Pa (500 mmHg). The sample was then evacuated at 2.67 Pa (20 mmHg) before N_2 adsorption[30,31]. The particle size distribution for both the materials was determined with Master sizer long bench particle size analyzer supplied by Malvern instruments, Canada. The instrument was provided with minimum He-Ne laser (633 nm wavelength), 18 beam diameter and 42 element composite solid state detector arrays. For this analysis, the particle size analyzer was fitted with lens in the size range of 300RF (Reserve Fourier) for measuring particles of size 0.05-900 µm. The percentage volume of particles was recorded with respect to particle diameter in µm. The FT-IR spectra for both, the immobilized lipases and the support materials, were obtained by KBr pelleting method using Perkin Elmer, FT-IR spectrum GX in the IR range of 400-4000 cm^{-1} (Ref. 32,33).

Reaction scheme

The overall reaction scheme for the esterification of total free fatty acids which were extracted from the hydrolysis of tuna fish oil, with lauryl alcohol in presence of immobilized PCL and TLL has been given in Reaction Scheme 1 to form DHA rich lauryl esters selectively. The reaction scheme explains the formation of lauryl ester by the breakdown of lipase-substrate complex in the presence of stoichiometric amount of alcohol.

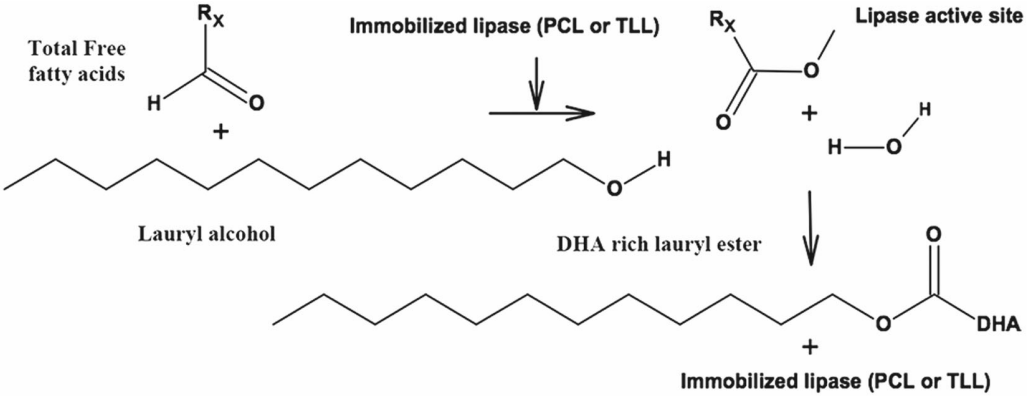
Results and Discussion

Characterization of immobilized lipases

The immobilized lipases such as PCL and TLL along with their support material immobead-150 were characterized for their physiochemical properties using various standard analytical instrumental methods such as Brunauer Emmett Teller (BET) and particle size distribution for surface properties and Fourier Transform-Infrared Spectroscopy (FT-IR) for the presence of functional groups.

BET analysis

The characterization of surface properties for both lipases and their support material have been performed with BET method and results are reported in Table 1. The immobilization of PCL and TLL on inert immobead-150 retains high activity and stability of these lipases. Their wide pores and large internal volume correspond to their activity and the ability to influence rate of reaction. The covalent bonding of immobead-150 with their respective lipases allow changes in the lipase conformation from close/inactive to open/active and provides the opening of the hydrophobic active sites towards the organic phase at the interface to minimize the dissolution of lipase in aqueous phase. Therefore, these lipases show better binding of substrate to the available enzymes active sites[34,35].



Reaction Scheme 1 — Esterification of free fatty acids with immobilized PCL or TLL in presence of lauryl alcohol

Table 1 — Characterization of immobilized lipases with BET method

S.No.	BET characteristics	Immobilized lipases		
		PCL	TLL	Immobead-150
1	BET surface area, S_g (m^2/g)	205.1±0.9	144.4±0.5	228.8 ±0.5
2	Pore diameter (A°)	100.1	98.1	89.2
3	Single point pore volume, V_v (cc/g)	0.56	0.35	0.5
4	Bulk density, ρ(g/cm^3)	1.8	2.9	2
5	Average pore dia., $D_p = 4. V_v/S_g$ (cm)	99.5	96.9	91.4
6	Porosity, $\varepsilon = V_v.\rho$	1.01	1.02	1

Particle size distribution analysis

The particle size distribution analysis was performed for immobilized lipases and their support as particle size has important role for explaining immobilization and enzyme activity[36]. The particle size distribution of immobead-150 was found in the range of 0-50 µm particle diameter. The trend in the variation of percentage volume with respect to particle diameter (µm) for immobilized PCL, TLL and immobead-150 has been compared in Figure 1. It was found that immobilized PCL and TLL enzymes show single narrow peak with average particle size of 300 and 500 µm respectively for a volume of 21-27% particles according to the Fig. 1. The support material immobead-150 has non-uniform distribution of particles in the wide range of 100-900 µm.

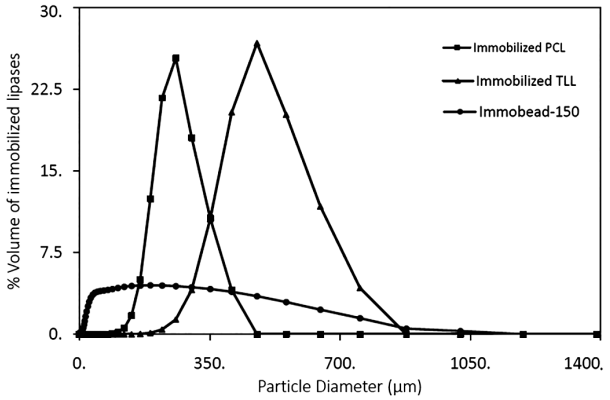


Fig. 1 — Particle size distribution for immobilized lipases and their support

FT-IR analysis

The FT-IR spectra of immobilized lipases and their support material were recorded separately as shown in Fig. 2. According to the spectra shown, the broad peak in the range of 3600-3200 cm^{-1} represents the free hydroxyl group of carboxylic acid. The strong sharp peak at 1700 cm^{-1} corresponds to carbonyl group of the carboxylic acid functional group. The medium peak at 2900 cm^{-1} is the characteristic peak for stretching vibration while strong peak at 1500 cm^{-1} is for the bending vibration of free amine present in the enzyme. According to the Figure, the small peaks ranging from 1200-600 cm^{-1} indicate the finger print region characteristic for immobilization support material immobead-150.

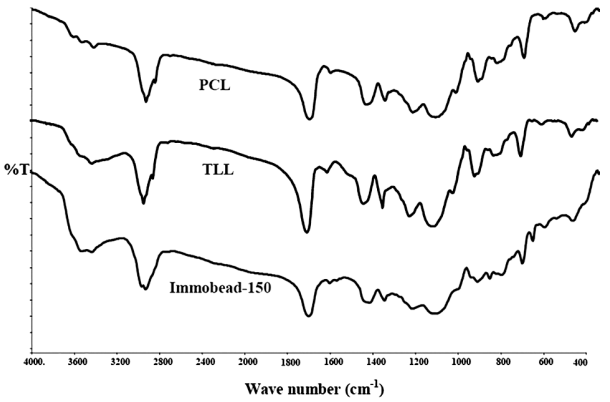


Fig. 2 — FT-IR spectrum of immobilized lipases and support

Parameter optimization

The optimization study of various reaction conditions have been carried out to maximize the production of DHA rich ester. The parameters optimized were type of lipases, type of alcohol, alcohol amount, *p*H, buffer amount, temperature and speed of agitation for the selective esterification of tuna free fatty acids with immobilized lipase and iso-octane solvent. The details of the studies and findings are given below.

Effect of different type of immobilized lipases

Four different immobilized lipases have been investigated for their ability to selectively perform the esterification of DHA from tuna free fatty acids. The screening of immobilized MML, PCL, TLL and ROL was performed at reaction conditions such as *p*H 7.5 and temperature 40°C using 1:1 (wt/wt) solvent to oil ratio, 2.0 g of butanol with 600 rpm speed of agitation. The 250 units (U) of immobilized PCL, TLL, ROL and MML enzymes corresponding to 0.27, 0.04, 0.68 and 2.3 g, were taken for the study respectively. The trend of variation for tuna free fatty

acid ester formation with various immobilized lipases is shown in Fig. 3 at time t = 0 and after esterification reaction performed for 6 and 24 h. According to the results reported, maximum DHA esters of 40.7 and 18.9 wt% were formed with immobilized PCL and TLL respectively than other lipases used in 6 h. Therefore, in the present study immobilized PCL has been chosen to increase the esterification of DHA than other fatty acids selectively. Mbatia *et al.*[37] have also reported immobilized PCL and TLL lipases for enrichment of DHA and EPA by selective esterification of free fatty acids or by hydrolysis of ethyl esters of fatty acids from oil.

Type of alcohol

The formation of DHA esters with selective esterification of TFFAs was studied with various alcohols such as ethanol (CH_3CH_2OH), methanol (CH_3OH), 1-propanol ($CH_3CH_2CH_2OH$), 2-propanol ($CH_3CHOHCH_3$), butanol ($CH_3CH_2CH_2CH_2OH$), hexanol ($CH_3(CH_2)_5OH$), heptanol ($CH_3(CH_2)_6OH$), ocatnol ($CH_3(CH_2)_7OH$), benzyl alcohol ($C_6H_5CH_2OH$), furfuryl alcohol (C_5H_6O) and lauryl alcohol ($C_{12}H_{26}O$). According to the results given in Table 2, although a maximum 48.3% esterification, was observed with butanol which corresponds to only 42.8 wt% DHA. On the other hand, lauryl alcohol (LA) was found to give maximum production of DHA esters i.e. 88.9 wt % (110.9µmoles/ml), corresponding to 33.7% esterification in 6 h. Therefore, lauryl alcohol was chosen to increase the DHA in esters selectively with immobilized PCL in the present study. This indicates that as the LA is a long carbon chain alcohol than other alcohols used in the study, it is able to bind with the long chain fatty acids specially DHA to form its esters. The results clearly indicate that the immobilized PCL can only bind with three types of fatty acids including $C_{18:0}$, $C_{22:6}$ and $C_{24:1}$ in the presence of lauryl alcohol, among which the maximum preference has been seen for DHA ($C_{22:6}$) in esters i.e. 88.9 wt%. Therefore, it is concluded from the present study that both lipase and alcohol are responsible for selectively enhancing the separation of DHA. Triantafyllou *et al.*[38] have used lauryl alcohol for the esterification of ethyl caprate in the presence of di-isopropyl ether with *Candida antarctica* lipase both in free and immobilized form to increase the activity of lipase. Gandhi *et al.*[39] have reported thermo deactivation of lipozyme (immobilized *Mucor miehei*) by exposure to butanol, whereas higher reaction rate was achieved in biphasic system with lauryl alcohol for the esterification of lauric acid.

Effect of alcohol amount

The effect of LA amount on formation of DHA esters from TFFAs has been studied to optimize its amount for the esterification reaction. The study has been conducted at different amounts of LA. The trend of variation in esterification of tuna free fatty acids with immobilized PCL is shown in Fig. 4. The results indicate that maximum 37.4% esterification was achieved with 4.0 g of lauryl alcohol per g of TFFAs, corresponding to 138.5 µmoles of fatty acid esters formed/ml of reaction mixture. The trend in Fig. 4 reflects that the increased amount of LA for

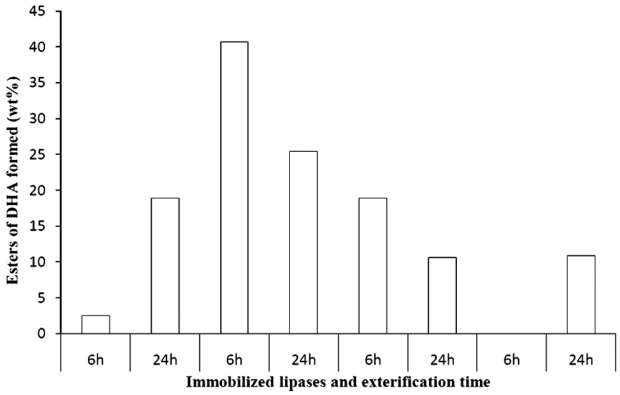


Fig. 3 — Effect of various immobilized lipases for the formation of DHA esters

Table 2 — Effect of different alcohols on activity of immobilized PCL for selective esterification of TFFAs

Type of alcohol	Fatty acid composition in esters (wt%)											%E
	C14:0	C16:0	C16:1	C18:0	C18:1	C18:2	C20:2	C20:5	C22:6	C24:1	Others	
Initial	0.8	13.8	10.9	15.2	6.4	4.6	6.5	0.6	35.6	1.8	3.8	0.0
Ethanol	10.5	-	33.2	24.4	0.33	14.1	-	0.2	-	8.6	8.7	17.4
2-Propanol	9.1	40.8	-	23.9	-	14.4	-	-	-	5.7	6.1	15.2
Butanol	-	-	-	11.1	1.4	6.2	9.4	-	42.8	2.8	26.3	48.3
Hexanol	-	-	-	-	11.3	-	6.8	0.9	10.7	-	70.3	20.7
Heptanol	-	-	-	-	-	-	-	19.4	4.4	9.4	66.8	22.4
Lauryl alcohol	-	-	-	-	-	3.8	-	-	88.9	7.3	0.0	33.7

esterification reaction was not able to produce higher amount of DHA esters. The possible reason seems to be the substrate inhibition of the reaction at higher amount of alcohol because alcohol is one of the substrate for esterification of fatty acids and according to the stoichiometry of the esterification reaction in presence of LA, equal molar ratio of fatty acids and alcohol is desirable to obtain best results. Syamsul *et al.*[40] have studied esterification at 1:2 molar ratios of palmitic acid to lauryl alcohol with commercial immobilized Novozym 435 *Candida antarctica* lipase-B and obtained a high purity yield (>90%) of lauryl palmitate at their optimized reaction conditions.

Effect of pH

The nature of acidic and basic environment in terms of *p*H (4.0 to 9.0) was evaluated for esterification reaction to produce DHA rich esters using immobilized PCL enzyme. The trend of variation in esterification of tuna free fatty acids with immobilized PCL is shown in Fig. 5. According to the results obtained, the maximum 35.6 % esterification of fatty acids is at *p*H 8.0 i.e. 131.9 µmoles of fatty acid esters formed in 6 h and 93.5% of maximum enzyme activity was noted at *p*H 7.0. It has been seen from Fig. 5, that the immobilized PCL is able to reflect optimum activity for esterification reaction in the basic (*p*H >7.0) conditions than the acidic (below *p*H < 7.0). The increase in *p*H from 4.0 to 8.0 shows an increasing trend in the enzyme activity but a sharp reduction in activity was observed at *p*H 9.0. The possible reason for this behavior of immobilized PCL is due to the presence of net positive charge on amino acid residues present in active sites of lipase at *p*H 8.0. Hence, positively charged active sites provide more binding affinity for negatively charged fatty acids residues to carry out esterification reaction optimally in basic biphasic solvent environment. Zhang *et al.*[41] have reported that the highest esterification of L-ascorbic acid occurred at neutral *p*H in case of lipase from *Thermomyces lanuginosus* immobilized on silica gel. Dosanjh and Kaur [42] have compared the physiochemical properties of *Bacillus* lipase in free and immobilized form and reported *p*H 8.0 as optimum condition for esterification.

Effect of buffer amount

The important aspect of an enzyme catalyzed esterification reaction in biphasic environment is to maintain the constant *p*H conditions as confirmed in the previous section throughout the course of reaction for achieving optimum enzyme activity because enzymes are very fragile in nature and their protein folding structure can be easily degraded upon exposure to variable adverse *p*H conditions. Therefore, it is crucial to maintain stable *p*H environment using appropriate amount of buffer which prevents the change in *p*H of a reaction mixture. The effect of buffer amount has been studied to maintain *p*H 8.0 in the present case and the trend of variation in esterification of tuna free fatty acids with immobilized PCL is shown in Fig. 6. Results show that the optimized amount of buffer was 1.0 g per g of TFFAs to produce 152.4 µmoles of fatty acid esters, which led to 36.6% esterification in 6 h whereas 124.8 µmoles of fatty acid esters was formed with 2.0 g of buffer per g of TFFAs. This reflects that esterification reaction can be optimally carried out with 1.0-2.0 g of buffer per g of TFFAs as not much variation in the amount of fatty acid esters product was obtained within this range. Various studies have shown that lipases need a small amount of water for esterification reactions to retain their active three-dimensional conformational state,

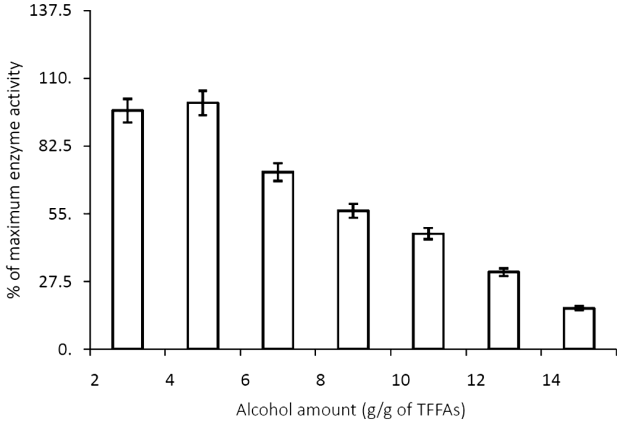


Fig. 4 — Effect of alcohol amount on activity of immobilized PCL

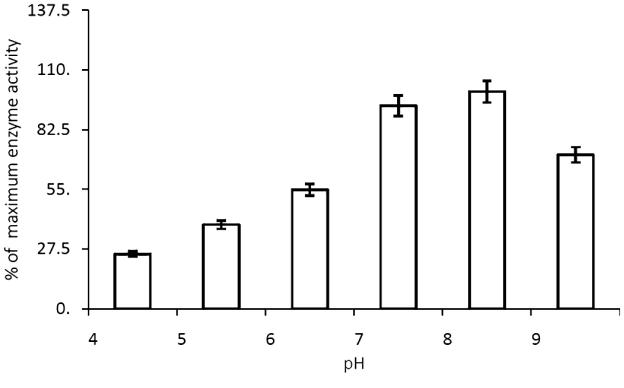


Fig. 5 — Effect of *p*H on activity of immobilized PCL

even when enzyme is covalently bound to support. Water can also limit the solubility of substrate around the enzyme even in the form of buffer[43,44].

Effect of temperature

In order to account the structural integrity and stability of an immobilized lipase with respect to temperature variation in the reaction conditions[45], the effect of temperature change on the esterification of free fatty acids into ester form was studied with immobilized PCL. From this study, it was concluded that maximum activity was observed at 50°C in terms of maximum 36.8% esterification corresponding to 153.3 µmoles fatty acid esters formed per mL of reaction mixture which was reduced to 96.7 µmoles/mL at 60°C in 6 h. The trend of variation in esterification of tuna free fatty acids with immobilized PCL is shown in Fig. 7. This reflects that under the present reaction conditions, the immobilized PCL remained stable only up to 50°C and then, its denaturation started causing drastic reduction in activity at higher temperature. The conversion decreased after 60°C probably caused by the vibration and movement of the enzyme molecule, which would affect the hydrogen bonds and other bonds in the lipase structure [46]. Mbatia *et al.*[37] have also used 50°C to carry out esterification reaction of PUFA extracted from Nile perch fish oil using immobilized PCL and TLL lipases.

Effect of agitation speed

The effect of agitation speed was studied for esterification reaction catalyzed by immobilized PCL enzyme to reduce the influence of bulk mass transfer resistances. For a reaction catalyzed in the presence of immobilized enzyme, it is significant to eliminate the effect of mass transfer resistances by providing desired contact between substrate and immobilized enzyme through optimum speed of agitation. The results indicate that a speed of 800 rpm was sufficient to maximize formation of fatty acid esters i.e. 159.6 µmoles/mL corresponding to 38.3% esterification in 6 h. The trend of variation in esterification of tuna free fatty acids with immobilized PCL is shown in Fig. 8. The increasing order of product formation was observed with increase in speed up to 800 rpm, due to reduction in the resistances generated by mass transfer. On the other hand, a sharp reduction in enzyme activity was found when further speed was increased from 800 to 1000 rpm due to the denaturation of immobilized PCL in terms of structure and activity.

Activation energy for esterification of TFFAs

The determination of activation energy for selective esterification of DHA with immobilized PCL and TLL was studied using Arrhenius plot as described in equation (1). The percentage conversion of tuna free fatty acids into esters of lauryl alcohol

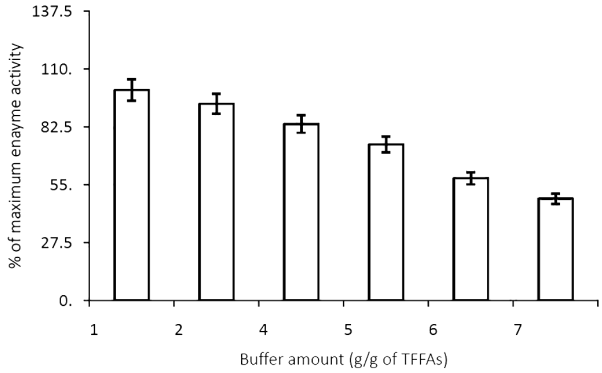


Fig. 6 — Effect of buffer amount on activity of immobilized PCL

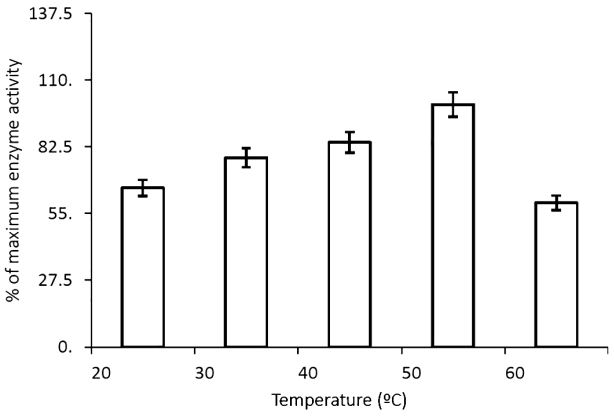


Fig. 7 — Effect of temperature on activity of immobilized PCL

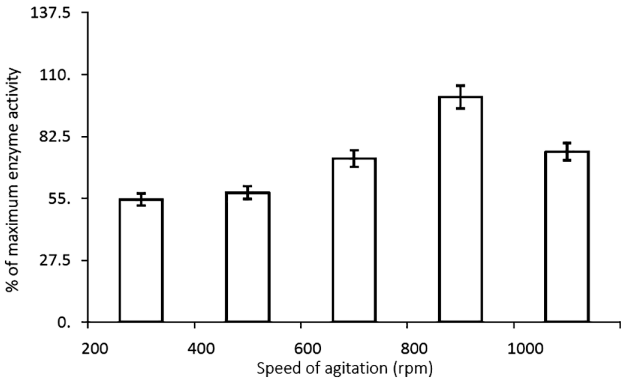


Fig. 8 — Effect of speed of agitation on activity of immobilized PCL

Table 3 — Effect of temperature on rate constant (k) with immobilized PCL and TLL

Temperature T (ºK)	Rate constant k (μmoles of DHA esters formed per ml)		1/T (ºK^{-1})	ln k	
	Immobilized PCL	Immobilized TLL		Immobilized PCL	Immobilized TLL
303	0.3674	0.2991	0.0033	-1.0	-1.35
313	0.44	0.3576	0.0032	-0.82	-1.03
323	0.6389	0.3762	0.0031	-0.45	-0.75

[Reaction Conditions: 0.5 g TFFAs, 0.5 g phosphate buffer of *p*H 8.0, 50°C, 800 rpm, 2.0 g lauryl alcohol, 1.0 g of iso-octane, immobilized PCL 0.54 and TLL 0.08 g and reaction time 6 h]

was plotted with time (up to 6 h) at varying temperatures such as 30, 40 and 50°C for immobilized PCL and TLL respectively.

The rate of the reaction increases with the temperature and facilitates more collision between reactants and enzyme. Therefore, reactants bind and product releases easily from the enzyme and this causes the reduction in the activation energy of the enzymatic system. The linearized form of Arrhenius law was used to calculate the value of activation energy. Therefore, ln k was plotted with respect to the reciprocal of temperature as given in Fig. 9 with data given in Table 3. The Arrhenius plot indicates that the slope of the linear equation is equivalent to the value of (-E/R) and intercept gives the pre-exponential factor (A_0). Hence, the activation energy (E) equivalent to 22.4 ($E_{imm.,PCL}$ and 24.9 ($E_{imm.,TLL}$) kJ/mol were determined from Figure 9 for immobilized PCL and TLL enzymes, respectively. The magnitude of activation energy (E) is a mathematical measure to represent the ease of reaction in presence of catalyst. The summary of results for activation energy study is given in Table 4.

In the present case, the amount of activation energy required by immobilized PCL and TLL enzymes ($E_{imm.,PCL}<E_{imm.,TLL}$) indicates that the esterification with immobilized PCL is comparatively more favorable than with immobilized TLL. For this system, to the best of our knowledge, no study has been reported earlier for estimation of activation energy. However, there are some studies reported on determination of activation energy for lipase catalyzed reactions such as alcoholysis of methyl propionate and 1-propanol with immobilized *Candida antarctica* lipase B in an organic liquid medium[47], esterification of lauric acid and 1-propanol with immobilized *Rhizomucor miehei* lipase[48] and esterification of lauric acid and lauryl alcohol with immobilized Porcine pancreatic lipase[49].

Table 4 — Activation energy (E) for selective esterification of TFFAs with immobilized PCL and TLL

Lipase	Activation energy, E (kJ/mol)	Pre-exponential factor, A_0
Immobilized PCL	22.8	31.12 x 10^2
Immobilized TLL	24.9	52.01 x 10^2

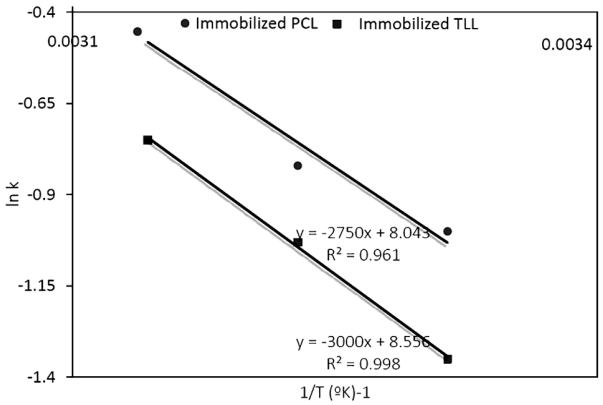


Fig. 9 — Arrhenius plot for selective esterification of TFFAs with immobilized PCL and TLL enzymes

Reusability study with immobilized PCL and TLL for esterification of TFFAs

In the enzymatic treatment processes, the high cost of the enzyme is a crucial parameter that should be controlled to make the process economically feasible and therefore, the cost of the process can be reduced by repeated use of the enzyme, even at comparatively low activity.

The reusability study for immobilized PCL and TLL enzymes were performed at optimized reaction conditions with 0.5 g of tuna free fatty acids and 0.53 and 0.08 g of immobilized PCL and TLL respectively. The reusability experiments were conducted up to five runs and reaction mixture was placed in the orbital shaker at 800 rpm for 6 h. After every run, the immobilized lipases were isolated by filtration, washed three times with iso-octane solvent (10 mL), distilled water (10 ml) and dried at room temperature. The results for reusability study are given in Table 5.

Table 5 — Reusability of immobilized PCL and TLL for esterification of TFFAs

	%E		Fatty acids in FFAs (wt%)		Fatty acid esters formed (μmoles/ml)		% Activity retention of immobilized lipases	
	PCL	TLL	PCL	TLL	PCL	TLL	PCL	TLL
1	47.4	39.7	52.6	60.3	197.5	165.4	100	100
2	43.6	35.2	56.4	64.8	181.7	146.7	91.9	88.7
3	38.9	28.5	61.1	71.5	162.1	118.8	82.1	71.8
4	29.2	19.8	70.8	80.2	121.7	82.5	61.6	49.8
5	28.9	18.9	71.1	81.1	120.4	78.8	60.9	47.6

[Reaction Conditions: 0.5 g TFFAs, 0.5 g phosphate buffer of *p*H 8.0, 50°C, 800 rpm, 2.0 g lauryl alcohol, 1.0 g iso-octane, immobilized PCL 250 U (0.53 g) and immobilized TLL 250 U (0.08 g) and reaction time 6 h]

The trend in the variation of percentage activity retention of immobilized PCL and TLL is shown in Fig. 10 for the esterification of tuna free fatty acids. The esterification of tuna free fatty acids with the repeated use of lipases was found to decrease from 43.6 to 38.9% and 35.2 to 28.5% after second and third runs for immobilized PCL and TLL, respectively (Table 5). Additionally, activity retention was calculated for repeated runs of immobilized PCL and TLL, which indicates that activity of lipase, has been reduced to 29.2 and 19.8% after fourth recycling. It indicates that immobilized lipases retains ability to react with the substrate even up to four repeated runs but the activity was found to decrease drastically. The reduction in activity is a clear indication of partial blockage of active sites present in lipase causing lesser conversion. To confirm this, the BET analysis was performed for the recycled immobilized PCL and TLL and compared with the fresh immobilized lipases. The results indicate that after fourth recycle of the immobilized PCL and TLL, the BET surface area was reduced to 115.4 and 69.7 m^2/g in comparison with 225.1 and 144.4 m^2/g for fresh immobilized PCL and TLL respectively.

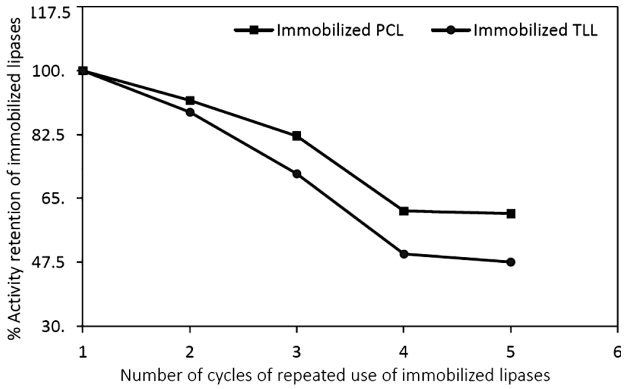


Fig. 10 — Reusability study with immobilized PCL and TLL for esterification of TFFAs with lauryl alcohol

Composition of fatty acids in esters after esterification of TFFAs with immobilized PCL

Esterification of tuna free fatty acids with immobilized PCL was carried out up to 56 h. The fatty acid esters of lauryl alcohol obtained after esterification at different time of reaction were separated from the reaction mixture standard procedure reported by Wanasundara and Shahidi[28] and Senanayake and Shahidi[29] with some modifications. The composition of the fatty acid esters obtained after 56 h of esterification reaction of tuna free fatty acids is given in Table 6. From the table, it can be seen that the DHA content of fatty acid esters obtained after esterification was maximum at 93.8 wt% after reaction time of 16 h when the extent of esterification was 32.9%. On the contrary, after 48 h the extent of esterification reached up to 57.5% and at this stage, the DHA content of fatty acid esters was found to be around 70.5 wt%. Higher concentration of DHA (93.8 wt%) in fatty acid esters of lauryl alcohol after the esterification for 16 h compared to that after esterification for 48 h clearly indicates that the immobilized PCL had preference for forming the DHA ester of lauryl alcohol over other fatty acids. Therefore, fatty acid esters formed after 16 h esterification was chosen for the isolation of DHA from the fatty acid esters of lauryl alcohol to study the synthesis of DHA rich glycerides. A batch of 5.0 g tuna free fatty acids was taken for the esterification of fatty acids with immobilized PCL. After esterification for 16 h and separation, the DHA rich (93.8 wt%) fatty acid esters obtained were 1.54 g. The DHA in free fatty acids was recovered and analyzed using HPLC with ELSD detector. The single peak shown in HPLC chromatogram at retention time of 14.6 min. indicates the presence of pure DHA rich FFAs which was isolated successfully. After the chemical breakdown and separation, 1.5 g (30 wt%) of DHA

Table 6 — Composition of fatty acid esters of lauryl alcohol with respect to time for esterification of TFFAs with immobilized PCL

Time (h)	%E	Fatty acids in FFAs (wt%)	Fatty acid composition of esters (wt %)		Composition of FFAs mixture (wt%)		Fatty acid esters formed (μmoles/mL)
			$C_{22:6}$	Other	$C_{22:6}$	Other	
0	0.0	100.0	-	-	35.6	64.4	0
0.1	6.5	93.5	49.6	50.4	34.3	65.7	28.9
0.3	7.7	92.3	54.2	45.8	33.5	66.5	33.7
0.5	10.0	90.0	56.8	43.2	32.3	67.7	44.7
0.8	11.5	88.5	60.9	39.1	31.0	69.0	51.3
1	14.4	85.6	63.7	36.3	28.9	71.1	64.3
2	20.9	79.1	70.4	29.6	24.9	75.1	77.9
4	23.0	77.0	77.5	22.5	18.0	82.0	98.3
6	25.7	74.3	88.6	11.4	7.6	92.4	114.9
8	31.6	68.4	91.6	8.4	5.4	94.6	140.9
16	32.9	67.1	93.8	6.2	7.9	92.1	147.3
20	38.9	61.1	86.8	13.2	14.3	85.7	173.9
24	48.0	52.0	78.5	21.5	14.5	50.7	212.6
32	51.3	48.7	74.7	25.3	15.7	47.0	228.9
40	55.8	44.2	73.3	26.7	19.1	44.5	248.9
48	57.5	42.5	71.9	28.1	19.8	42.7	256.9
56	56.9	43.1	70.5	29.5	18.1	41.9	253.9

[Reaction Conditions: 5.0 g TFFAs, 5.0 g phosphate buffer of pH 8.0, 50ºC, 800 rpm, 10.0 g iso-octane solvent, 20.0 g lauryl alcohol and 5.4 g (250 U) of immobilized PCL]

rich free fatty acids was collected. The DHA recovered was 84.3% with respect to initial DHA content of tuna free fatty acids (35.6 wt%).

Conclusion

The most suitable reaction conditions for selective esterification of TFFAs with immobilized PCL, were established with 1:1 (wt/wt) isooctane and 250 U of enzyme as TFFAs amount: 0.25g, , speed of agitation: 800 rpm, pH 8.0, TFFAs to buffer ratio: 1:1 (wt/wt), and lauryl, TFFAs to temperature: 50°C alcohol ratio: 1:4 (wt/wt) for immobilized PCL.

The activation energies for esterification of TFFAs were found to be 22.8 and 24.9 kJ/mol with immobilized PCL and TLL, respectively, in the temperature range of 30-50°C, indicating that the esterification with immobilized PCL is comparatively more favorable than with immobilized TLL. In 48 h, maximum 57.5% esterification was obtained with immobilized PCL. But maximum DHA content in esters was found to be 93.8 wt% after 16 h at 32.9% esterification. The high concentration of DHA (93.8 wt%) in esters of lauryl alcohol after esterification for 16 h compared to that after 48 h clearly indicated that the immobilized PCL had preference for forming the DHA ester of lauryl alcohol over other fatty acids. Therefore, fatty acid esters formed after esterification for 16 h was chosen for the isolation of DHA from the fatty acids for the further study. The reusability of immobilized PCL and TLL for selective esterification of TFFAs with lauryl alcohol was also studied and activity retentions of 61.6 and 49.8% respectively, were found after four time repeated use of the immobilized lipases.

Acknowledgement

Ms. Sharma would like to thank you to the funding provided by Canada Research Chair (CRC) Program and by Natural Sciences and Engineering Council (NSERC) of Canada in terms of student assistance-ship to perform her research work in the Department of Chemical & Biological Engineering, College of Engineering, University of Saskatchewan, Saskatoon, Canada.

Nomenclature

PCL	*Pseudomonascepacia* lipase	TFFAs	Tuna free fatty acids
TLL	*Thermomyceslanugi nosus*lipase	DHA	Docosahexaenoi c acid
MML	*Mucormiehei* lipase	EPA	Ecosahexaenoic acid
ROL	*Rhizopusoryzae* lipase	FAMEs	Fatty acid methyl esters
% E	Percentage esterification	GC	Gas chromatography
HPLC	High performance liquid chromatography	LA	Lauryl alcohol

References

1 Kuo C H, Huang C Y, Chen J W, David Wang H M, Shieh C J, *Applied Biochemistry and Biotechnology*, 192 (2020) 517.

2 Lee M C, Jiang X, J. Brenna T & Abbaspourrad A, *Food Funct*., 9 (2018), 5598.

3 Zhang Z, Liu F, Ma X, Huang H & Wang Y, *J. Agri. Food Chem*., 66 (2018) 218.

4 Cervera M A R, Barriga V G, Romero J, Rojas R & Aran S L*, Foods*, 9 (2020) 233.

5 Stoneham T R, Kuhn D D, Taylor D P, Neilson A P, Smith S A & Gatlin D, *PLoS ONE*, 13 (2018), e0194241.

6 Oliver L, Dietrich T, Marañón I, Villarán M C & Barrio R J, *Resources*, (9) 2020, 148.

7 Patel A, Rova U, Christakopoulos P & Matsakas L*, Sci. of the Total Environment*, 736 (2020) 139691.

8 Rosabal G D, Rodríguez A, Contreras E, Viedma J O, Muñoz M, Trigo M, Aubourg S P & Espinos A, *Molecules*, 24 (2019) 1642.

9 Bispo P, Batista I, Bernardino R J & Bandarra N M, *Appl Biochem Biotechnol*., 172 (2014) 1866.

10 Kjerstad M, Larssen W E & Midtbo L K, *Heliyon*, 6 (2020) e05239.

11 Gámez Meza N, Noriega Rodríguez J A, Medina Juárez L A, Ortega García J, Monroy Rivera J, Toro Vázquez F J, García H S & Angulo Guerrero O, *Food Res Int*, 36 (2003) 721.

12 Ferna´ndez Lorente G, Pizarro C, Lo´pez Vela D, Betancor L, Carrascosa A V, Pessela B & Guisan J M, *J Am Oil Chem Soc*, 88 (2011) 819.

13 Ramírez Fajardo A, Esteban Cerdán L, Robles Medina A, MuñíoMartínez M M, Hita Peña E & Molina Grima E, *J Am Oil Chem Soc*, 83 (2006) 215.

14 Nunes P A, Cabral P P, Guillen M, Valero F & Dias S F, *Biochem Eng J*, 67 (2012) 265.

15 Valverde L M, Moreno P A G, Quevedo A R, Pena E H, Callejon L E, Cerdan E M & Garima A R M, *J Am Oil Chem Soc*, 89 (2012) 1633.

16 Pizarro C, Branes M C, Markovits A, Fernandez Lorente G, Guisan J M, Chamy R & Wilson L, *J Mol Catal B: Enzym*, 78 (2012) 111.

17 Brenneis R & Baeck B, *Biotechnol Lett*, 34 (2012) 1459.

18 Giua L, Cossignanai L, Simonetti M S, Lombardi G & Blasi F, *Eur Food Res Technol*, 235 (2012) 53.

19 Shimada Y, Sugihar A & Tominaga Y, *J Biosci Bioeng*, 91 (2001) 529.

20 Bhandari K, Chaurasia S P, Dalai A K & Gupta A, *J Am Oil Chem Soc*, 90 (2013) 1637.

21 Bhandari K, Chaurasia S P, Dalai A K, Gupta A &Singh K, *J Mol Catal B*, 94 (2013) 104.

22 Shimada Y, Sugihar A, Nakano H, Kuramoto T, Nagao T, Gemba M & Tominaga Y, *J Am Oil Chem Soc*, 74 (1997) 97.

23 Sharma A, Chaurasia S P & Dalai A K, *Canadian J Chem Eng*, 92 (2014) 344.

24 Hoshino T, Yamane T & Shimizu S, *Agric Biol Chem*, 54 (1990) 1459.

25 Shimada Y, Maruyama K, Sugihara A, Baba T, Komemushi S, Moriyama S & Tominaga Y, *JAOCS*, 75 (1998) 1565.

26 Shimada Y, Watanabe Y, Sugihara A, Baba T, Ooguri T, Moriyama Z S, Terai T & Tominaga Y, *Lipase J Biosci Bioeng*, 92 (2001) 19.

27 Chang Q L, Lee C H & Parkin K L, *Enzym Microb Technol*, 25 (1999) 290.

28 Wanasundara U N & Shahidi F, *Food Chem*, 65 (1999) 41.

29 Senanayake S & Shahidi F, *Food Chem*, 65 (1999) 41.

30 Ferdous N, Bakhshi N, Dalai A K & Adjaye J, *Appl Catal B: Environ*, 72 (2007) 118.

31 Surisetty V R, Dalai A K & Kozinski J, *Appl Catal A: General*, 385 (2010) 153.

32 Naik S, Goud V V, Rout P K & Dalai A K, *Biores Technol*, 101 (2010) 7605.

33 Naik S, Goud V V, Rout P K, Jacobson K & Dalai A K, *Renew Energy*, 35 (2010) 1624.

34 Hanefeld U, Gardossi L & Magner E, *Chem Soc Rev*, 38 (2009) 453.

35 Miranda M, Urioste D, Souza L T A, Mendes A A & De Castro H F, *Enzym Res*, (2011) 1.

36 Li Y, Gao F, Weia W, Qua J B, Guang-Hui M & Wei Qing Z, *J Mol Catal B: Enzym*, (2010) 182.

37 Mbatia B, Mattiasson B, Mulaa F & Adlercreutz P, *J Lipid Sci Technol*, 113 (2011) 717.

38 Triantafyllou A O, Wehtje E, Adlercreutz P & Mattiasson B, *Biotechnol Bioeng*, 54 (1997) 67.

39 Gandhi N N, Sawant S B, Joshi J B & Mukesh D, *Enzyme Microb Technol*, 17(1995) 373.

40 Syamsul K M W, Salina M R, Sitti S O, Hanina M N, Basyaruddin M A R & Jusoff K, *World Appl Sci J*, 11 (2010) 401.

41 Zhang D H, Li Y Q, Li C, Lv Y Dosanjh N S & Kaur J, *Appl Biotechnol Biochem*, 36 (2002) 7.

42 Dosanjh N S & Kaur J, *Appl Biotechnol Biochem*, 36 (2002) 7.

43 Tanaha A & Sonomoto K, *Chem Technol*, 20 (1990) 112.

44 Zaks A & Klibanov A M, *J Biol Chem*, 263 (1988) 3194.

45 Keng P S, Basri M, Abdul Rahman M B, Salleh A B & Ariff A B, *J Oleo Sci*, 54 (2005) 519.

46 Klinbanov A M, *Trends Biotechnol*, 15 (1997) 97.

47 Graber M, Dubouch M P B, Sousa N, Lamare S & Legoy M D, *Biochim Biophys Acta*, 1645 (2002) 56.

48 Zoumpanioti M, Parmaklis P, De Maria P D, Stamatis H, Sinisterra J V & Xenakis A, *Biotechnol Lett*, 30 (2008) 1627.

49 Gogoi S, Hazarika S, Rao P G & Dutta N N, *Biocatal Biotransform*, 24 (2006) 343.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LSP.2021.3090673, IEEE Signal Processing Letters

Local Moment driven PVO based Reversible Data Hiding

Neeraj Kumar, Rajeev Kumar, and Roberto Caldelli, *Senior Member, IEEE*

***Abstract*— Pixel-value-ordering (PVO) is one of the most widely used reversible data hiding (RDH) framework which efficiently utilizes smooth pixels of the cover image to provide high-fidelity stego-image but with limited embedding capacity. This letter proposes an RDH scheme based on local moment driven pixel value ordering (LM-PVO) which further enhances the effectiveness of smooth pixel's utilization by dividing the fixed-size blocks into two groups. Pixels of each group are sub-divided into two sub-groups based on the local moment of the block so that correlation among the pixels of each sub-group is enhanced. Thus doing, the pixels of each sub-group are grouped based on their intensity values instead of their position as in the conventional PVO-based schemes; this enables information hider to embed a higher amount of secret data while also enhancing the stego-image quality. Experimental results also validate the superiority of the proposed scheme over the existing PVO-based RDH schemes.**

***Index Terms*— Local Moment driven PVO, LM-PVO, pairwise embedding, reversible data hiding.**

I. INTRODUCTION

REVERSIBLE data hiding (RDH) is a method to securely embed a secret message into a cover/host image without permanently distorting it. Applications like military and medical image processing which require at the receiving end both high-quality host image and secret message decoding have been using the RDH methods quite extensively [1, 2]. Due to this, there has been quite significant growth in the development of high-fidelity RDH methods in the last decade. One such method, popularly known as prediction error expansion (PEE) is given by Sachnev et al. [1] in 2009. The PEE strategy leverages predictor (e.g., rhombus predictor) to predict reference pixel value and calculates prediction error with respect to the actual pixel value to create prediction error histogram (PEH).

This letter is submitted to the journal on 28 April 2021 for the review.

N. Kumar is with the Department of Electronics & Communication Engineering, Jamia Milia Islamia, Delhi, India (e-mail: neeraj.mohiwal@gmail.com).

R Kumar is with the Department of Computer Science & Engineering, Delhi Technological University, Delhi, India (e-mail: rajeevgargnsit@gmail.com).

R. Caldelli is associated with the National Inter-University Consortium for Telecommunications (CNIT), Parma, Italy and Mercatorum University, Rome, Italy (e-mail: roberto.caldelli@unifi.it).

Next, it expands the peak prediction errors (based on PEH) and shifts others (to ensure reversibility) which in turn modifies the pixel value to embed the secret message bits. Thus, the PEE method successfully increases the EC but introduces some distortion. In 2013, Ou et al. [3] unfolded the two-dimensional (2D) scenario of PEE known as pairwise PEE. The pairwise PEE made a slight but impactful change in the mapping of PEH by converting it into a 2D PEH with modified mapping. The 2D PEH mapping is generated by considering prediction errors in pairs corresponding to a pair of pixels by assuming that the spatially correlated prediction errors have a correlation as well as the adjacent pixels. Thus, the pairwise PEE boosts overall EC and also enhances marked-image quality. The pairwise embedding in the PEE scheme is further researched in [4-6] to again improve its performance.

In addition to the pairwise PEE, a noteworthy development in the high-fidelity RDH domain was introduced by Li et al. [7] that proposed the popularly known pixel value ordering (PVO). The PVO method partitions the cover image into uniform size blocks and then sorts the pixels in ascending order. The Li et al.'s method uses PEH to embed the secret message bits into the maximum and minimum valued pixels of each block. The method provides very good quality marked-image but with limited EC. To increase the EC and further enhance the image quality, various PVO-based RDH methods have been introduced which are reviewed in [9]. However, there is still a gap in the exploitation of correlation in block formation. To address this issue, the proposed letter introduces local moment driven PVO (LM-PVO) based reversible data hiding which uses local moment (or mean) to identify pixels that can form a smoother block. The proposed scheme additionally makes use of enhanced pairwise embedding to efficiently embed the secret message. Thus, the scheme is able to enhance the image quality and also increase the EC. The contribution of the LM-PVO based RDH scheme can be summarized as follows:

- LM-PVO exploits the correlations of nearby pixels and constitutes highly correlated pixel groups based on local moments, unlike the traditional schemes which group pixels based on their location.
- LM-PVO treats the image as a chessboard which renders independent context for the hider to apply both the pixel-wise and block-wise embedding strategies.
- The proposed scheme adaptively selects the embedding strategies for embedding the secret data based on the

1070-9908 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:27:34 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LSP.2021.3090673, IEEE Signal Processing Letters

number of pixels so that optimal embedding performance can be achieved.

- An enriched marked image with a significant improvement in PSNR along with embedding capacity is achieved.

The rest of the letter is organized as follows. Section II presents the related works. Section III introduces the proposed LM-PVO scheme. The experimental results are analyzed in Section IV and finally, the letter is concluded in Section V.

II. Related Works

In this section, some of the most popular and related pixel value ordering-based reversible data hiding schemes are briefly reviewed. One of the prominent early improvements in the PVO-RDH [7] was introduced by Peng et. al. [8] in 2013 in the form of Improved PVO (IPVO). The IPVO scheme makes use of two '0' and '1' prediction errors (instead of just '1' as in the case of PVO [7]) which are usually the peak bins in the PEH for hiding the secret message bits (SMB) by employing new difference calculation and histogram modification strategy. However, the middle pixels of each block remained non-utilized for embedding. To overcome this limitation, Kumar et al. [10] introduced a new PVO-RDH scheme based on block extension strategy. This scheme extends the block size (by 2×1 pixels) in case the noise level (i.e., the difference between the second and the third pixel) of the sorted four-pixel block is zero. The scheme embeds the secret data in the original four-pixel block using [8] and uses the PEE strategy with a new sorting-based predictor for embedding the secret message bits into the additional two pixels. Thus, the number of unused pixels such as the middle pixels in [7-8] are reduced which in turn increases the EC without noticeably deteriorating the image quality.

Wu et al. [11] introduced an adaptive pairwise PVO-based RDH scheme (based on [3]) to exploit the correlation among the prediction errors. The adaptive pairwise PVO considers the two prediction errors as a pair and modifies them for data embedding based on adaptive two-dimensional histogram modification. Thus, the scheme provides an improvement in terms of both image quality and EC. Zhang et al. [13] extend the work of pairwise PVO by considering both the pixel value orders and pixel locations in the predictor. The location-based PVO additionally introduces a self-learning mechanism to adaptively select the 2D mapping according to the image content. Thus, the scheme provides the good quality marked image. In another work of pairwise embedding, He et al. [12] discussed a multi-pass pixel value ordering (MPPVO) method which focuses to completely exploit the redundancy among the elements of each block for prediction, so that the least distortion can be caused to a marked image. Thus, these pairwise PVO-based RDH schemes provide high fidelity marked image (HFMI), however, their EC was limited. To overcome the limitation of EC while providing HFMI, Kumar et al. [14] discussed an enhanced pairwise pixel value ordering-based RDH scheme that uses two layers of embedding. Kumar et al. introduce a rhombus context based double layer embedding strategy for reversible data hiding by treating the host image as a chessboard. It forms image blocks of size 1×3 pixels by scanning the image in snake scan order. The blocks are categorized into smooth and complex categories based on the local complexity (calculated using rhombus mean) of each pixel. In the case of smooth blocks, the secret data is embedded into two layers: in the first layer, the embedding is done using enhanced pairwise IPVO while in the second it is done using a novel error-recovery based pairwise embedding strategy. Both the layers are designed in such a way that the pixels modified in the first layer are most probably recovered in the second layer. However, in the worst case, the second layer may again modify the pixels to cause more damage which in turn drastically reduces the image quality. In the case of complex blocks, pixel-wise embedding is performed. Overall, Kumar et al.'s technique is able to significantly increase the embedding capacity while the quality of the marked-image is a bit deteriorated.

Based on the above analysis, it is observed that existing PVO-based RDH schemes have not efficiently exploited the image correlation to form highly correlated image blocks as only the pixel locations are being used for block formation. To avoid this problem, the proposed letter suggests a local moment driven RDH scheme that effectively aggregates correlated pixels in the same group. This is done by sub-dividing the pixels of a block into groups based on their local moments and their location. Therefore, LM-PVO forms groups of highly correlated pixels, unlike the traditional PVO-based RDH schemes where highly correlated pixels may be present in different groups and vice-versa.

III. Proposed LM-PVO Scheme

This section presents the proposed LM-PVO scheme. Firstly, the image scanning and sub-groups generation strategy are introduced, then the rationale behind the generation of highly correlated pixel sub-groups based on local moment is presented. Finally, the adaptive embedding scheme based on the number of pixels in the sub-groups is discussed.

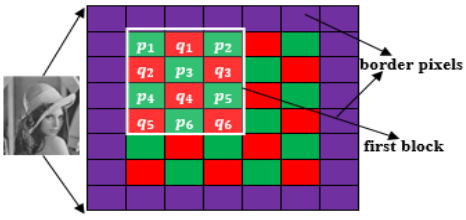


Fig. 1. Representation of block of 4×3 pixels

A. Image Scanning and sub-groups generation

Initially, the proposed LM-PVO scheme uniformly partitions the host image into non-overlapping blocks of size 4×3 pixels excluding the image border pixels (which are shown in violet color in Fig. 1). Next, the pixels of each block are arranged in two groups by treating the image as a chessboard. This way, the first group say $F_G = \{p_1, p_2, p_3, p_4, p_5, p_6\}$ and the second group say $S_G = \{q_1, q_2, q_3, q_4, q_5, q_6\}$ are obtained, where all the green pixels are grouped into the first group (F_G) and all the red pixels into the second one (S_G) as shown in Fig. 1. Each

1070-9908 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:27:34 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LSP.2021.3090673, IEEE Signal Processing Letters

group is further divided into two highly correlated sub-groups by separating the pixels according to the local moment (μ) and the moment of moments (Ω) which are calculated using Equations (1) and (2), respectively, defined as follows:

$$\mu = \left[\frac{1}{4}(q_1 + q_2 + q_3 + q_4)\right] \quad (1)$$

$$\Omega = \left[\frac{1}{M}\sum_{m=1}^{M}\mu_m\right] \quad (2)$$

where M represents the total number of pixels in the group (which are usually 6 in case of 4×3 sized block) and q_1, q_2, q_3, and q_4 are the surrounding pixels for the reference pixel say p_3, as shown in Fig. 1, for which the local moment (μ) is calculated. Next, the pixels whose local moment (μ_m) is greater than or equals to the mean of moments (Ω) of the group are grouped into the first sub-group F_G otherwise, they are placed into the second sub-group S_G. Doing so, it can be observed that the two sub-groups can have an unequal number of pixels but are highly correlated. Thus, the formation of sub-groups based on the mean of moments places the high correlated pixels in the same sub-group, unlike the traditional schemes which just group pixels based on their location in the image.

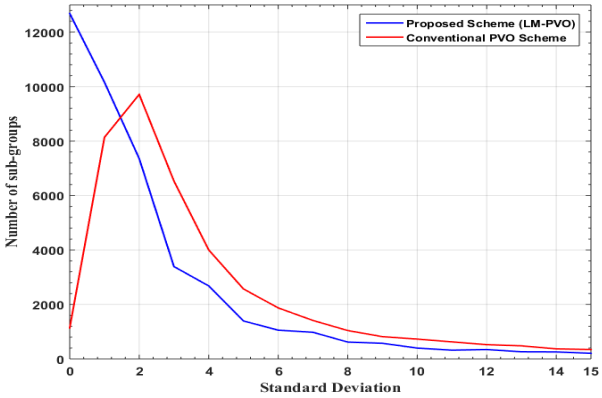
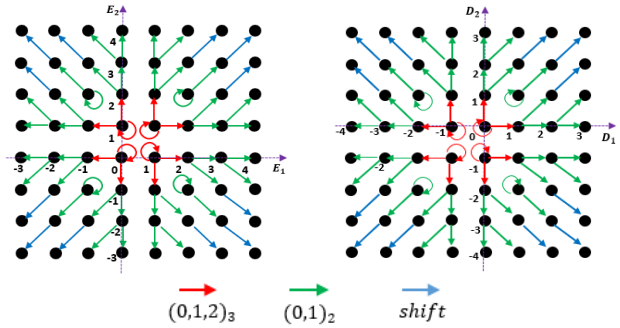


Fig. 2 Histogram of standard deviation of pixel sub-groups for Lena.

B. Rationale behind generation of highly correlated pixel sub-groups based on Local Moment

The local moment has been used for forming highly correlated sub-groups as it has been proved in [1, 3, 14] that any pixel can be best represented/predicted by its local context i.e., moment/mean of surrounding pixels. Additionally, the grouping of pixels of each block into two groups has allowed the user a free hand in the identification and modification of the pixels without worrying about reversibility. To validate the argument that local moment-based pixel distribution generates highly correlated sub-groups, a histogram-based comparison using an exemplary Lena image is presented in Fig.2. The plotted histograms are the histograms of the standard deviation of pixel groups for the proposed scheme and the conventional PVO (C-PVO) scheme (by considering block sizes of 2×3 pixels); for C-PVO, blocks, used for embedding, are considered. Fig. 2 shows that the proposed scheme has a higher peak for low values of standard deviation in comparison to the C-PVO that has the peak around 2. Thus, it is validated that the pixel distribution in the sub-groups generated by the LM-PVO has less variation, in other words, the sub-groups generated by LM-PVO have a superior correlation with respect to the conventional PVO scheme.



(a) EPPVO mapping strategy (b) EPPEE mapping strategy

Fig. 3: 2D mapping strategy

C. Adaptive Embedding Scheme based on number of pixels

The proposed RDH scheme adaptively selects embedding strategy based on the number of pixels in a sub-group. Initially, the embeddability of each sub-group is checked using local mean divergence (lmd) which is computed using Eq. (3), defined as follows, against a user-defined threshold.

$$lmd = max(\mu_1, \ldots \mu_N) - min(\mu_1, \ldots \mu_N) \quad (3)$$

where, $\mu_1, \ldots \mu_N$ are the local moments for pixels $p_1, \ldots p_N$ of the sub-group. For example, if the user-defined threshold (thr_1) is 5 and the computed lmd is 3 for the sub-group, then ($lmd < thr_1$) the sub-group is an embeddable one otherwise is not. Therefore, only the sub-groups having the maximum difference between their pixel values, indicated by local mean divergence (lmd) which is determined using their corresponding local moments, below the user-defined threshold (thr_1) are used to embed the secret message.

After this, if the sub-group contains more than two pixels, then enhanced pairwise PVO (EPPVO) strategy is used for embedding the message bits. For applying the EPPVO, first, the pixels of the sub-group are arranged in ascending order ($p_{\pi(1)}, \ldots\ p_{\pi(N)}$) where $p_{\pi(1)} \leq \cdots \leq p_{\pi(N)}$, $\pi(i) < \pi(j)$ if $p_{\pi(i)} = p_{\pi(j)}$ and $i < j$. Next, two prediction errors are calculated such as $E_1 = p_s - p_t$ and $E_2 = p_u - p_v$ where $s = min(\pi(1), \pi(2))$, $t = max(\pi(1), \pi(2))$, $u = min(\pi(N-1), \pi(N))$ and $v = max(\pi(N-1), \pi(N))$ [8,14]. Now, the value of minimum valued pixel ($p_{\pi(1)}$) is either decreased by '1' or remains unchanged and the value of maximum valued pixels ($p_{\pi(N)}$) is either increased by '1' or remains unchanged based on the bit value of the secret message. The modification of the prediction errors is done using the EPPVO strategy which expands the error pair (0, 0) to only three pairs (0, 0), (0, -1), and (-1, 0), and thus only $\log_2 3$ bits are embedded into this pair. Moreover, the pair (-1, -1) is expanded to itself and (-2, -2) to embed 1-bit data in EPPVO, while this pair can be just shifted to (-2, -2) in the conventional pairwise I-PVO. The other pairs like (1,1), (0,1), and (1,0) are also similarly expanded as the pair (0,0). The complete illustrative expansion and shifting mechanism for EPPVO are shown in Fig. 3(a). On the contrary, if the sub-group has only two pixels, then, first of all, the

1070-9908 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:27:34 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LSP.2021.3090673, IEEE Signal Processing Letters

embeddability of each pixel is determined by measuring the local complexity (lc) of each pixel which is calculated using Eq. (4) and compared with a user-defined threshold (thr_2).

$$lc = \sqrt{(q_1 - q_2)^2 + (q_2 - q_3)^2 + (q_3 - q_4)^2 + (q_4 - q_1)^2} \quad (4)$$

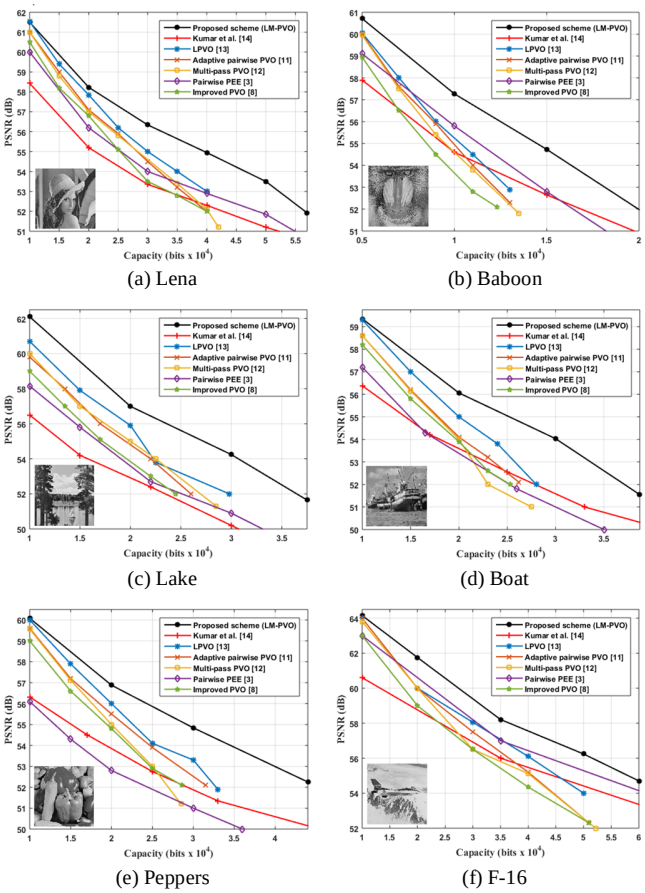
where q_1, q_2, q_3, and q_4 are the surrounding pixels for the reference pixel (e.g. p_3), as shown in Fig. 1. For example, if the user-defined threshold (thr_2) is 8 and lc is 12 for the pixel p_3, then ($lc < thr_2$) the pixel is an embeddable one otherwise is not. Therefore, only the pixels having the local complexity (lc) less than the user-defined threshold (thr_2) are used to embed the secret data.

Now, if both the pixels of the sub-group are determined as embeddable ones, then, in this case, enhanced pairwise PEE (EPPEE) scheme is used for embedding the secret message. For applying the EPPEE, two prediction differences (errors) are calculated such as $D_1 = p_1\text{-}\mu_1$ and $D_2 = p_2\text{-}\mu_2$, where μ_1 and μ_2 are the local moments of the pixels p_1 and p_2, respectively. Now, the modification of the prediction differences is done using EPPEE mapping strategy [3] as shown in Fig. 3(b). As per the EPPEE, the error pairs such as (0, 0) are expanded to only three possible pairs (0, 0), (0, 1), and (1, 0) to embed $\log_2 3$ bits, while discarding expansion from (0, 0) to (1, 1) to minimize distortion but at the cost of EC. However, it additionally exploits prediction error pair (1,1) to embed one-bit data by expanding (1,1) to either (1,1) or (2,2). Similarly, the error pairs in the other three quadrants are expanded/shifted. Finally, in case the sub-group has only one pixel then no embedding is carried out as it has been observed that the pixel generally belongs to the outlier category of the sub-groups. Thus, all the pixels of the group F_G are processed to embed the message bits. Similarly, the other group of pixels S_G of the block is processed for embedding. Such a procedure goes on in a block-wise manner to embed all the message bit-stream.

To blindly recover the host image and extract the hidden message on the receiver side, some auxiliary information which includes location map, last embeddable block, thresholds, etc., is also embedded into the host image along with the secret data. Due to the scarcity of space, the discussion regarding the location map and auxiliary information generation and management is omitted as these steps are similar to Kumar et al.'s scheme [14]. The extraction process just follows the embedding steps in reverse order.

IV. Experimental Results and Discussion

This section presents the experimental results and their comparative analysis for the proposed LM-PVO scheme with some of the most popular and recently developed schemes such as Improved PVO [8], Multi-pass PVO [12], LPVO [13], adaptive pairwise PVO [11], Kumar et al. [14]. Furthermore, it is also compared with the pairwise PEE technique [3] as the LM-PVO makes the adaptive selection of embedding strategies from pairwise PVO, pairwise PEE, and conventional PEE. The experimental results in terms of PSNR and EC are taken on widely used six standard gray-scale images of 512×512 pixels, taken from the USC-SIPI database [15]. Fig.4 illustrates the performance comparison in terms of PSNR versus EC of the proposed scheme (LM-PVO) with all the other existing schemes. The experimental results of Fig.4 are taken by setting a user-defined threshold thr_1 at 5 for Lena & F-16 while it has been set at 10 for other images. As far as the user-defined threshold thr_2 is concerned, its value ranges from 3 to 25 to get the optimal performance. Similar iterative optimization has been performed for all the compared methods. From Fig. 4, it is evident that LM-PVO achieves better performance in terms of PSNR vs EC than all the existing schemes. The main reason behind its performance is the generation of high-correlated sub-groups which makes the embedding task easier. Furthermore, the adaptive selection of different embedding strategies also helps in enhancing the embedding performance.



(a) Lena (b) Baboon

(c) Lake (d) Boat

(e) Peppers (f) F-16

Fig.4: PSNR versus EC: comparison of the proposed LM-PVO with the methods in [3, 8, 11, 12, 13, 14]

V. Conclusions

In this letter, a high-fidelity RDH scheme based on local moment driven pixel value ordering (LM-PVO) has been introduced. The LM-PVO scheme sub-divides the image blocks to form highly correlated pixel groups using local moments of the pixels and then makes an adaptive selection of embedding strategies based on the number of pixels in each group. Experimental results validate that the proposed scheme grants an improvement in both PSNR and embedding capacity in comparison to existing high-fidelity RDH schemes. Future works will be dedicated to further enhance group correlation.

1070-9908 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:27:34 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LSP.2021.3090673, IEEE Signal Processing Letters

REFERENCES

[1] V. Sachnev, H. J. Kim, J. Nam, S. Suresh and Y. Q. Shi, "Reversible watermarking algorithm using sorting and prediction", IEEE Trans. Circuits Syst. Video Technol., vol. 19, no. 7, pp. 989-999, Jul. 2009.

[2] J. Tian, "Reversible data embedding using a difference expansion", IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 8, pp. 890-896, Aug. 2003.

[3] B. Ou, X. Li, Y. Zhao, R. Ni and Y. Q. Shi, "Pairwise prediction-error expansion for efficient reversible data hiding", IEEE Trans. Image Process., vol. 22, no. 12, pp. 5010-5021, Dec. 2013.

[4] B. Ou, X. Li, W. Zhang and Y. Zhao, "Improving Pairwise PEE via Hybrid-Dimensional Histogram Generation and Adaptive Mapping Selection", IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 7, pp. 2176-2190, 2019.

[5] I. Dragoi, D. Coltuc and I. Caciula, "Horizontal pairwise reversible watermarking," Proc. *EUSIPCO*, pp. 56-60, 2015.

[6] I. C. Dragoi, D. Coltuc and H. G. Coanda, "Adaptive pairwise reversible watermarking with horizontal grouping", Proc. *ISSCS*, pp. 1-4, 2017.

[7] X. Li, J. Li, B. Li and B. Yang, "High-fidelity reversible data hiding scheme based on pixel-value-ordering and prediction-error expansion", Signal Process., vol. 93, no. 1, pp. 198-205, 2013.

[8] F. Peng, X. Li and B. Yang, "Improved PVO-based reversible data hiding", Digit. Signal Process., vol. 25, pp. 255-265, 2014.

[9] G. Kaur, S. Singh, R. Rani and R. Kumar, "A Comprehensive Study of Reversible Data Hiding (RDH) Schemes Based on Pixel Value Ordering (PVO)", Archives of Computational Methods in Engineering, 2020. https://doi.org/10.1007/s11831-020-09512-3

[10] R. Kumar, D. Kim, S. Lim and K. Jung, "High-Fidelity Reversible Data Hiding Using Block Extension Strategy," Proc. *ITC-CSCC*, pp. 1-4, 2019.

[11] H. Wu, X. Li, Y. Zhao and R. Ni, "Improved reversible data hiding based on PVO and adaptive pairwise embedding", Journal of Real-Time Image Processing, vol. 16, pp. 685–695, 2019.

[12] W. He, G, Xiong, S. Weng, Z. Cai and Y. Wang, "Reversible data hiding using multi-pass pixel value-ordering and pairwise prediction-error expansion", Information Sciences, vol. 467, pp. 784-799, 2018.

[13] T. Zhang, X. Li, W. Qi and Z. Guo, "Location-Based PVO and Adaptive Pairwise Modification for Efficient Reversible Data Hiding," IEEE Transactions on Information Forensics and Security, vol. 15, pp. 2306-2319, 2020.

[14] R. Kumar and K.-H. Jung, "Enhanced pairwise IPVO-based reversible data hiding scheme using rhombus context", Information Sciences, vol. 536, pp. 101-119, 2020.

[15] The USC-SIPI Image Database, [online] Available: http://sipi.usc.edu/database.

1070-9908 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:27:34 UTC from IEEE Xplore. Restrictions apply.

Lower Fin Modulation Analysis for a Novel 5nm Top Bottom Gated Junctionless FinFET for improved performance

Lakshya Gangwani
Department of Applied Physics
Delhi Technological University
New Delhi, India
lakshyagangwani1@gmail.com

Ruchika Chakravarti
Department of Applied Physics
Delhi Technological University
New Delhi, India
ruchika.chakravarti1@gmail.com

Rishu Chaujar
Department of Applied Physics
Delhi Technological University
New Delhi, India
chaujar.rishu@dtu.ac.in

2021 Devices for Integrated Circuit (DevIC) | 978-1-7281-9955-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/DevIC50843.2021.9455834

***Abstract*— In this paper, lower fin modulation is performed for a Novel 5nm Top Bottom Gated Junctionless Fin shaped Field Effect Transistor (FinFET). The analysis is done for fin lengths varying from 3nm to 12nm with and without source, drain contacts present at the bottom fin and compared with the Conventional Junctionless FinFET structure. The Novel Top Bottom Gated Junctionless FinFET with 3nm Lower Fin length and contacts on bottom fin shows 197 times improvement in switching ratio as compared to the Conventional Junctionless FinFET at the same gate length. The results reveal that the device shows enhancement in terms of switching ratio and allows better gate controllability when stacked against the Conventional Junctionless FinFET for various fin lengths**.

Keywords— Top Bottom Gated, FinFET, Fin Modulation, Switching Ratio, Junctionless

I. INTRODUCTION

Technology and along with it our way of life is ever evolving. Innovation is happening in all fields of research with greater importance being given to portability, efficiency and speedy operation. Keeping in line with this trend the semiconductor and solid state technology field has also experienced development with greater impetus for designing compact and low power consumption devices. Major players in the industry are already adapting and catering to these needs [1, 2]. These innovations entail scaling down of devices with more integration and lower power consumption. But during the process of scaling down we face certain obstacles. These obstacles take the shape of Short Channel Effects (SCEs) which reflect inefficient device performance and delays. These include leakage current, Hot Carrier Effects, Velocity Saturation and Drain Induced Barrier Lowering (DIBL) [3]. New improvements in technologies like SOI and FinFET are being made to address these effects. In this work we have performed lower fin dimension modulation for a Novel 5nm Top Bottom Gated (TBG) Junctionless (JL) Fin shaped Field Effect Transistor (FinFET). The analysis is done for fin lengths varying from 3nm to 12nm with and without source, drain contacts present at the bottom fin and compared with the Conventional Junctionless (JL) FinFET structure to record the improvements in switching ratio, TGF, transconductance and transfer characteristics thus reducing Short Channel Effects (SCEs).

II. DEVICE STRUCTURE AND METHODOLOGY

Fig. 1 depicts the Novel 5nm Top Bottom Gated Junctionless FinFET device having 5nm lower fin with source, drain contacts. The gates on both sides of the fins enable the device to perform better in terms of output current and switching ratio. It also helps in enhancing the control of gate over the device as compared to the Conventional Junctionless (JL) FinFET.

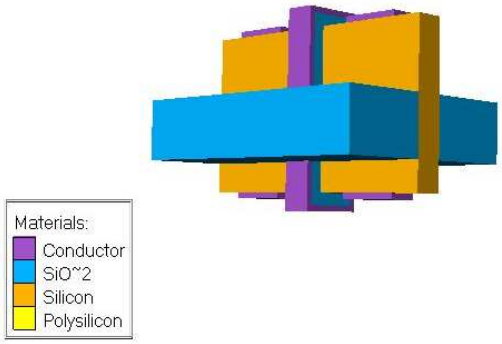


Fig. 1 Novel Top Bottom Gated Junctionless FinFET Structure with 5nm Lower Fin

In contrast to the Novel Top Bottom Gated (TBG) Junctionless (JL) FinFET, the Conventional Junctionless (JL) FinFET has gate above the substrate only due to the absence of the lower fin. As a result, it has less gate controllability over the device.

The device structures of both the Novel Top Bottom Gated (TBG) Junctionless (JL) and Conventional Junctionless (JL) FinFET have 5nm gate length, 35nm device length and 20nm source, drain length. For the case of Novel TBG JL FinFET, the device height varies from 31nm for 3nm lower fin to a maximum of 40nm for 12nm lower fin. Whereas on the other hand, the Conventional JL FinFET has fixed device height of 25nm due to absence of the lower fin. The additional lower fin in Novel TBG JL FinFET with contacts helps in enhancing the overall device performance in comparison to the Conventional JL FinFET.

Both the Novel TBG JL FinFET and Conventional JL FinFET are n-type devices with 5.0×10^{18} (cm^{-3}) uniform doping for source, drain and channel regions.

The Novel TBG JL FinFET has a gate length of 5nm and the lower fin dimension is varied from 3nm to 12nm. Also source, drain contacts are placed on the lower fin for all the device lengths and their results are analyzed and compared with the Conventional JL FinFET. The device has 2nm thickness of gate oxide for both the Conventional JL FinFET and Novel TBG JL FinFET [3].

Larger fin heights implies more gate control over the channel as the gate contact is covering significantly larger area over the channel for the same fin width. Thus it improves the output current and transconductance. But on the other hand, larger gate area leads to more surface scattering and thus leads to delays in switching. Due to this, smaller fin heights

978-1-7281-9955-9/21/$31.00 ©2021 IEEE
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:57:42 UTC from IEEE Xplore. Restrictions apply.

have better switching ratio as compared to larger ones. Also placing the contacts on the lower fin provides another path for the carriers to flow, thereby improving the output current considerably as compared to the structures without the lower fin contacts.

The three-dimensional (3D) device simulations for all the structures having different fin lengths is done on Silvaco Atlas TCAD Simulator using DeckBuild and TonyPlot tools [4]. The device simulation in Silvaco Atlas is divided into five main steps, videlicet the Specification of the Device Structure, Specification of Material Models, Selecting Numerical Models, Specification of the Solution and Analyzing the Results. Structure Specification basically involves defining the mesh grid for the device by using the mesh points and spacing, defining various regions in device, electrode and doping definition for the device. Material Model Specification involves defining the materials to be used in the device and choosing the physical models for the device. Numerical Method Selection is used to decide the mathematical models to be used for device simulation [4]. The device temperature was kept at T=300K for the analysis on Silvaco TCAD. Solution Specification and Result Analysis uses the extract and solve for running the simulations and extracting the required parameters.

Table 1. STRUCTURAL PARAMETERS SPECIFICATIONS

Parameters	**Dimensions of Simulated Devices**	
	Novel TBG JL FinFET	***Conventional JL FinFET***
Gate Length	5nm	5nm
Upper Fin Height	12nm	12nm
Lower Fin Height	3nm to 12nm	0nm
Fin Width	5nm	5nm
Effective Oxide Thickness	2nm	2nm
Substrate Width	5nm	5nm
Source,Drain Length	20nm	20nm

III. SIMULATED RESULTS AND DISCUSSIONS

Several device performance parameters namely Transconductance (g_m), Output Conductance (g_d), Transconductance Generation Factor (TGF) (g_m/I_d), Early Voltage (V_A) (I_d/g_d) are germane to the discussion while analyzing the Analog Performance of FinFET devices [5].

The ratio of ON state current I_{ON} and OFF state current I_{OFF} of a switch is a design parameter and is known as switching ratio. The ON current is recorded when the device allows current to flow through it and hence is 'ON' while the OFF current is ideally null [6]. But when we employ solid state technologies like FinFET to perform switching operations we do not obtain zero as OFF state current but rather it has a very small value and we get this value when $V_{gs} = 0$ Volts and $V_{ds} = V_{dd}$. As we scale down the device this small value can increase many folds and so becomes more important and is a reflection of Short Channel Effects (SCEs) and leakage in the device. Hence we try to minimize the OFF current which leads to high switching ratio and lower the losses.

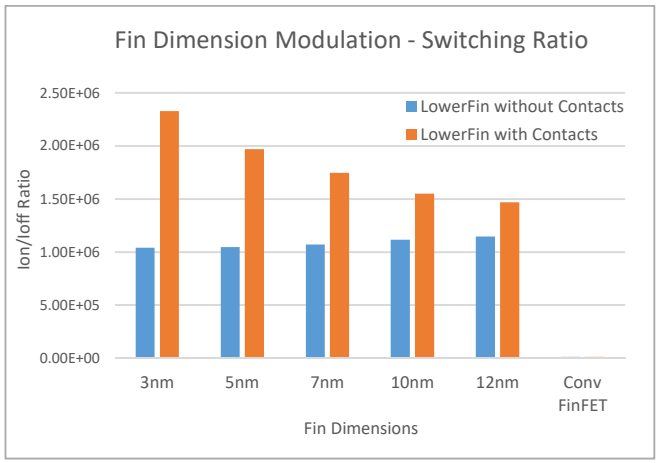


Fig. 2 Switching Ratio Variation with Lower Fin Dimensions for Novel TBG JL FinFET

Fig. 2 shows the switching ratio for different lower fin dimensions varying from 3nm to 12nm [7]. The ratios are given for both with contacts and without source, drain contacts placed on lower fin for the Novel Top Bottom Gated Junctionless FinFET. The trend can be observed that the switching ratio decreases as we increase the lower fin length for the device with contacts whereas the switching ratio increases with increase in lower fin length in the case of no contacts. Further it can be seen that the maximum switching ratio is obtained for the configuration of 3nm lower fin length with contacts. We can also observe that the switching ratio for all fin lengths of Novel Top Bottom Gated Junctionless FinFET is higher than Conventional JL FinFET. Thus indicating improvement.

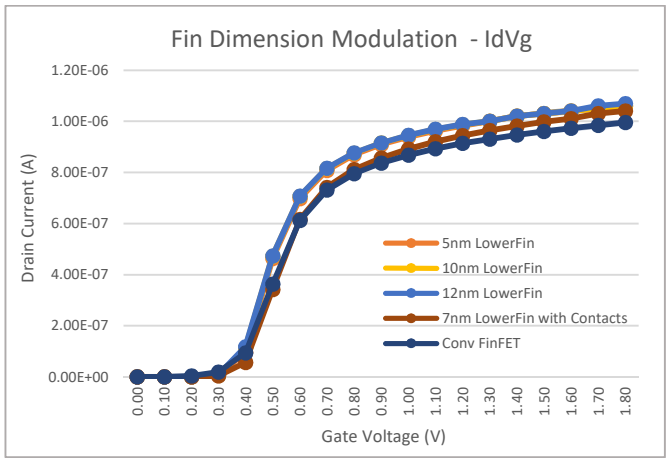


Fig. 3 Transfer Characteristics for Novel TBG JL FinFET for different lower fin lengths

Transfer Characteristics essentially shows the variation of drain current or the output current with the gate voltage (input voltage) [7]. **Fig. 3** depicts the $I_{ds} - V_{gs}$ curves for different lower fin lengths ranging from 3nm to 12nm with and without source, drain contacts placed on lower fin for the Novel Top Bottom Gated Junctionless FinFET. It is observed that 12nm lower fin without contacts shows increased output current for the same gate voltages as compared to the other fin lengths. All the fin lengths from 3nm to 12nm give better results in terms of output current as compared to the Conventional Junctionless FinFET.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:57:42 UTC from IEEE Xplore. Restrictions apply.

Thus 12nm lower fin Novel TBG JL FinFET structure without contacts gives best performance in terms of $I_{ds} - V_{gs}$ closely followed by 10nm lower fin Novel TBG JL FinFET structure. Higher drain current for 12nm lower fin depicts that gate has enhanced control over the device and V_{ds} effect on output is less, thus helping in improvement of device performance.

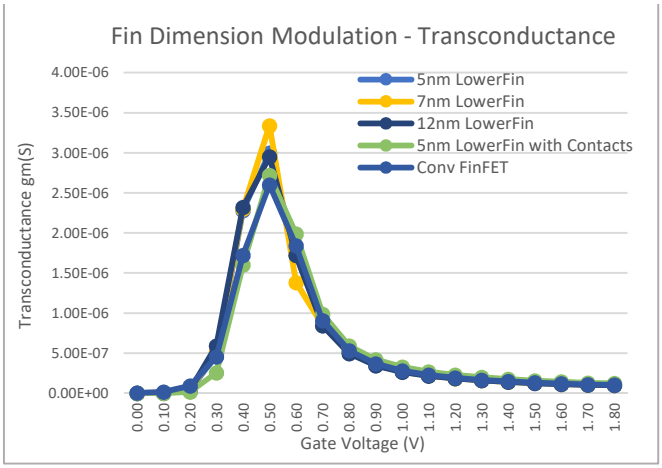


Fig. 4 Transconductance variation with fin dimension for Novel TBG JL FinFET

The linear gain of a device decides the input conductance (g_m or also known as transconductance). The transconductance is thus proportional to it. It depicts the fluctuations in output current with changes in input voltage V_{gs} under the condition where drain voltage V_{ds} is constant [8].
Transconductance is given by:

$$g_m = \frac{\partial I_d}{\partial V_g}$$

So we easily calculate g_m by the slope of $I_{ds} - V_{gs}$ plot [5]. **Fig. 4** shows the transconductance (g_m) for different lower fin lengths ranging from 5nm to 12nm with and without source, drain contacts placed on lower fin for the Novel Top Bottom Gated Junctionless FinFET with gate bias varying from 0 to 1.8 Volts. And the same is also plotted for Conventional Junctionless FinFET for comparison. We can see that we get highest transconductance value for 7nm lower fin without contacts as compared to the other fin lengths at the same gate voltage of 0.5 Volts. All the fin lengths from 5nm to 10nm give improved g_m values as compared to Conventional Junctionless FinFET.

Transconductance Generation Factor (TGF) is a measure of the efficiency of a device. It conveys how adequately the current is used to achieve a certain transconductance [8].
It is calculated as:

$$TGF = \frac{g_m}{I_d}$$

The merit of high TGF is the capacity for functioning of device at low voltages and therefore this leads to more efficient utilization and performance [5].
Fig. 5 shows the variation of TGF for different lower fin lengths of 5nm and 12nm with and without source, drain contacts placed on lower fin for the Novel Top Bottom Gated Junctionless FinFET. Again we have presented the same for Conventional Junctionless FinFET to compare performance.

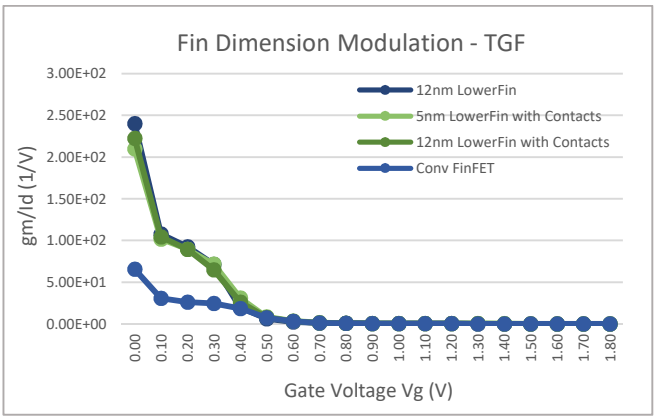


Fig. 5 Variation of Transconductance Generation Factor for different lower fin lengths for Novel TBG JL FinFET

We observe that TGF is highest at 0V gate supply and has maximum value 12nm lower fin length without contacts and shows quiet improvement over Conventional JL FinFET. Thus indicating attainment of superior parameters [7].

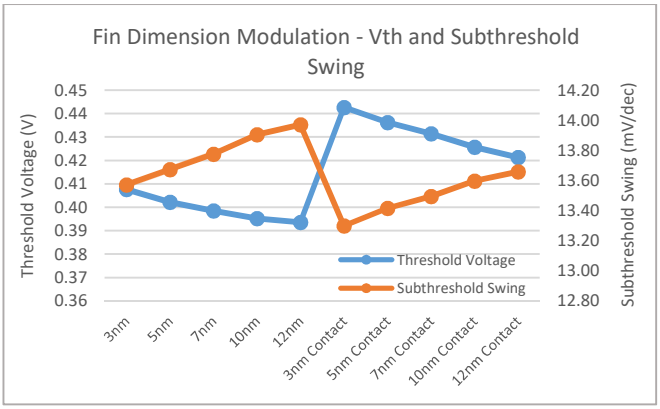


Fig. 6 Variation of Threshold Voltage and Subthreshold Swing for Novel TBG JL FinFET for varying lower fin lengths

V_{th} or the Threshold Voltage is the voltage required at the gate terminal to turn ON the device. At this voltage, the current in the device begins to appreciably [6].
Consequently Subthreshold Swing or also known as the Slope Factor is the inverse of the slope of log (I_d) vs. V_{gs} plot in the Subthreshold region. Conventional MOSFET have their Subthreshold Swing around 60mV/decade [7].
Fig. 6 shows the variation of threshold voltage and the Subthreshold Swing for different lower fin lengths with and without source, drain contacts for the Novel TBG JL FinFET structure.
It can be clearly inferred from the plot that threshold voltage decreases with increase in fin length and the lowest threshold voltage is achieved for 12nm lower fin without contact at V_{th} = 0.39V. Also threshold voltage for devices having lower fin with source, drain contacts is comparatively higher than their counterparts without the source, drain contacts. Thus lower fin without contacts leads to less threshold voltages, thereby low power consumption and device ON at lower voltages [9].
As for the Subthreshold Swing, it increases with increase in lower fin length and the lowest Subthreshold Swing is observed for the 3nm lower fin Novel TBG JL FinFET with contacts. The Subthreshold Swing for 3nm Lower Fin Novel TBG JL FinFET with contacts is S.S=13.29 mV/decade. It is

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:57:42 UTC from IEEE Xplore. Restrictions apply.

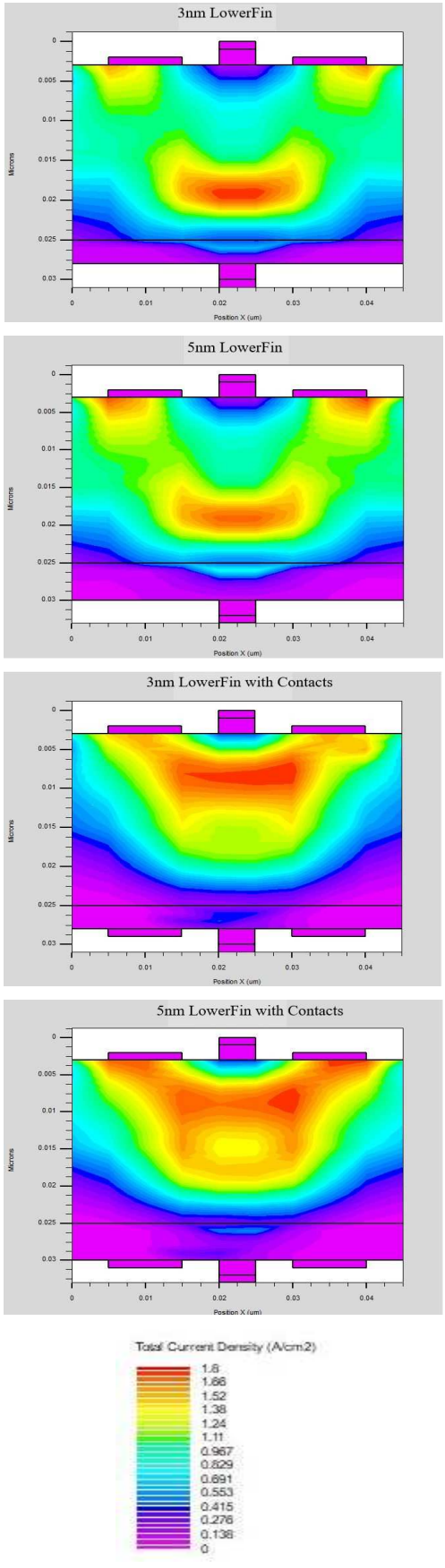


Fig. 7 Current Density Contours for Novel TBG JL FinFET for various lower fin dimensions

Also recognized that devices having lower fin with contacts have lower values for Subthreshold Swing as compared to their equivalent structures without contacts. Minimizing the Subthreshold Swing helps in reducing the leakage current up to a huge extent. The device has Subthreshold Swing less than 60mV/decade for all the fin lengths [3].

Fig. 7 shows the current density contours for the Novel Top Bottom Gated Junctionless FinFET at various lower fin lengths with and without source, drain contacts on lower fins. The current density primarily depicts the current flowing through the device per unit area. It is clearly visible from the contours that Novel TBG JL FinFET devices having lower fin contacts have higher current densities as compared to their complement structures without the contacts. This is due to the fact that contacts on lower fins provide another junction for the carriers to flow from and to, thus leading to increase in the current densities [10].

The 3nm Lower Fin Novel TBG JL FinFET with contacts shows the maximum current densities out of all the simulated devices, thus leading to improved switching ratio [11]. But higher current densities have tradeoffs as well. Higher current densities can lead to melting or heating of the insulating material and can also lead to electromigration.

IV. Conclusion

In this paper we have performed the lower fin modulation for a Novel 5nm Top Bottom Gated (TBG) Junctionless (JL) Fin shaped Field Effect Transistor (FinFET). The 3nm lower fin Novel TBG JL FinFET device showed considerably improved switching ratio (197 times over Conventional FinFET) and lowest Subthreshold Swing 13.29mV/decade is observed for the 3nm lower fin length as well. Whereas 7nm fin length without contacts had best I_dV_g characteristics and 12nm lower fin length without contacts has most improvement (2.65 times improvement over Conventional FinFET) for the Transconductance Generation Factor in comparison to other fin dimensions [12]. Thus the simulation results reveal that all the fin lengths present enhancement over the Conventional JL FinFET and the device can be utilized for high performance analog use [7].

References

[1] TSMC. Leading Edge Technology. http://www.tsmc.com/english/dedicate-Foundry/technology/16nm.htm. Accessed 3rd February 2021

[2] Semiconductor Engineering. IBM, Intel and TSMC Roll out FinFETs. http://semiengineering.com/ibm-intel-and-tsmc-roll-outfinfets. Accessed 3rd February 2021

[3] M. P. Kashyap and R. Chaujar, "Temperature investigation of a novel 3nm TF-Bulk FinFET for improved performance," 2020 IEEE 20th International Conference on Nanotechnology (IEEE-NANO), Montreal, QC, Canada, 2020, pp. 382-387, doi: 10.1109/NANO47656.2020.9183594.

[4] I. SILVACO, "ATLAS User's Manual," Santa Clara, CA, Ver, vol. 5, 2011.

[5] N. Kaur, M. Rattan & S.S Gill, "Design and optimization of novel shaped FinFET". *Arab J Sci Eng* 44, 3101–3116 (2019). https://doi.org/10.1007/s13369-018-3428-3

[6] S. Tripathi, N. Gupta, P. Tripathi & R. Chaujar. (2019). "GaN Silicon-on-Insulator (SOI) N-Channel FinFET for high-performance low power applications". 1-4. 10.1109/NMDC47361.2019.9084011.

[7] A. Kumar, N. Gupta & R. Chaujar, "Analysis of novel transparent gate recessed channel (TGRC) MOSFET for improved analog behaviour".

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:57:42 UTC from IEEE Xplore. Restrictions apply.

Microsyst Technol **22,** 2665–2671 (2016). https://doi.org/10.1007/s00542-015-2554-z

[8] M Singh, S Mishra, S S Mohanty, and G P Mishra. "Performance analysis of SOI MOSFET with rectangular recessed channel" Advances in Natural Sciences: Nanoscience and Nanotechnology, vol. 7, no. 1, 2016. doi:10.1088/2043-6262/7/1/015010

[9] A. Bughio, S. D. Guerrieri, F. Bonani, and G. Ghione, "Multi-Gate FinFET Mixer Variability Assessment through Physics-Based Simulation," IEEE Electron Device Letters, vol. 38, no. 8, pp. 1004-1007, 2017.

[10] S.Y. Kim, J. Seo, Y.J. Yoon, G.Yoo, Y.J. Kim, H. Eun et al. (2014). "Design and Analysis of Sub-10 nm Junctionless Fin-Shaped Field-Effect Transistors". Journal of Semiconductor Technology and Science. 14. 508-517. 10.5573/JSTS.2014.14.5.508.

[11] M. Tan & R. Ismail (2007). "Modeling of Nanoscale MOSFET Performance in the Velocity Saturation Region". Elektrika: Journal of Electrical Engineering.

[12] A. Kumar, B. Tiwari, S. Singh, M. M. Tripathi, and R. Chaujar, "Radiation Analysis of N-Channel TGRC-MOSFET: An X-Ray Dosimeter," IEEE Transactions on Electron Devices, vol. 65, no. 11, pp. 5014-5020, 2018.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:57:42 UTC from IEEE Xplore. Restrictions apply.

LRD-Net: A Lightweight Deep Network for Low-light Image Enhancement

Shivam Singhal
Department of Computer Science and Engineering
Delhi Technological University
New Delhi, India
shivam_2k17co323@dtu.ac.in

Srishti Nanduri
Department of Computer Science and Engineering
Delhi Technological University
New Delhi, India
srishti.nanduri@gmail.com

Yash Raghav
Department of Computer Science and Engineering
Delhi Technological University
New Delhi, India
1yashraghav@gmail.com

Anil Singh Parihar
Department of Computer Science and Engineering
Delhi Technological University
New Delhi, India
parihar.anil@gmail.com

2021 3rd International Conference on Signal Processing and Communication (ICPSC) | 978-1-6654-2864-4/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICSPC51351.2021.9451681

***Abstract*—The computer vision algorithms demonstrate superior performance in the real world when provided with high-quality image or video inputs. However, low-light or non-uniformly illuminated conditions result in low visibility in the captured images, thus demanding low-light image enhancement. Therefore, the paper presents a lightweight deep learning-based method, referred to as LRD-Net, for enhancing low-light images. The proposed approach initially estimates the illumination map using maximum of RGB color channels and then utilizes the retinex theory to derive the corresponding reflectance. Further, we propose a lightweight deep network for refining the estimated reflectance, which is capable of restoring the lost structural and textural details. Our proposed approach employs a lightweight structure yet demonstrates competitive performance to various state-of-the-art methods. The fast computation speed escalates the practical usage of our method without compromising on quality, as demonstrated by extensive experiments.**

***Index Terms*—Low-light image enhancement, reflectance, illumination, deep convolutional neural network.**

I. Introduction

There are significant advancements in the camera technology over the past decades leading to high-quality images. While modern cameras provide various options for capturing moving, distant or ultra-wide scenes, they often fail in capturing insufficiently illuminated scenes. Images clicked under poor or non-uniform illumination conditions generally have low visibility accompanied by noise and poor contrast. It affects the visual quality of images and also impacts the ever-growing applications of computer vision such as automated driving, face recognition, and so on. Thus, enhancing the low-light images becomes a critical task. A straightforward approach involves simply increasing the overall brightness of low-light images, but we rarely have uniformly illuminated images. In most low-light images, the darkness is non-uniform, thus losing the details present in already bright regions on direct brightness adjustment. Therefore, there is a need to focus on more sophisticated techniques that can restore the details in non-uniform illumination images while dealing with various color, contrast, and noise issues.

Some early researches include histogram equalization [1], [2] and gamma correction [3], [4], [5] based methods for enhancing low-light images. However, such techniques generally increase the overall contrast of image without taking the illumination factors in consideration. CLAHE [6] proposes a modification to HE by focusing on local contrast rather than global contrast, thus dealing with over-enhancement of bright regions. Some other approaches derived from retinex theory [7] include single-scale [8] and multi-scale retinex [9].

The growing interest in deep learning and its notable performance on various tasks in image processing, including dehazing [10], image-segmentation [11], etc. has led to the introduction of various deep networks [12], [13], [14] for enhancing low-light images. Msr-net [15] views the previously introduced MSR in terms of a convolutional neural network and maps the low-light images to normal-light images based on an end-to-end mapping. However, various deep learning methods based on encoder-decoder [16] or other similar networks consist of a large number of layers making such techniques computationally expensive. Thus, the paper focuses on building a lightweight deep learning based method for low-light image enhancement.

The main contributions of this study are described below:

- We propose a novel retinex theory based model, referred to as LRD-Net, for low-light image enhancement that works with little pre or post-processing.
- We leverage a lightweight deep network that directly learns to adjust the estimated reflectance based on the proposed multi-objective loss function.
- We conduct extensive experiments and compare the performance of proposed method with other best alternatives.

The remaining paper is arranged in the following manner. Section II presents some previous works that closely relate to our method. Section III has the details of the proposed ap-

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:30:05 UTC from IEEE Xplore. Restrictions apply.

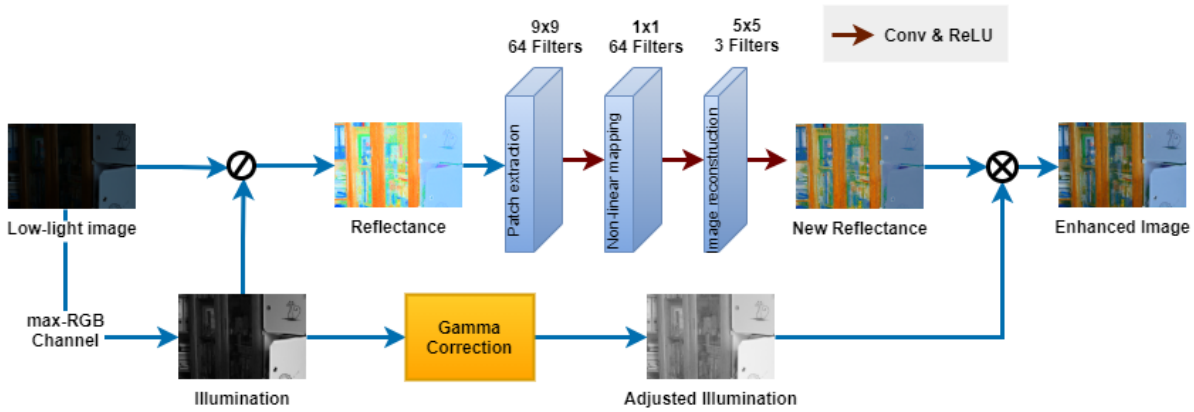


Fig. 1: The proposed framework of LRD-Net.

proach. Section IV comprises of experimental details followed by results. Eventually, section V concludes the proposed work.

II. Related Work

A wide variety of low-light image enhancement techniques have been introduced so far. Our proposed LRD-Net takes inspiration from the previously introduced works, and we briefly discuss some closely related approaches in this section.

The initial attempts for enhancing low-light images relate to histogram equalization [1], [2], [6] and gamma correction [3], [4], [5]. Some other early approaches [7], [8], [9] are based on retinex theory, and are discussed by various researchers [17], [18]. Such techniques consider the estimated reflectance component as the final enhanced image leaving behind the information carried by the illumination component. SRIE [19] introduced a weighted variational model that estimates both reflectance and illumination simultaneously. SRIE provides noise suppression and maintains color constancy to give visually pleasant output. NPEA [20] uses a bright-pass filter for image decomposition and preserves the details by using bi-log transformation while adjusting the estimated illumination. Guo *et al.* [21] introduced a structure prior for estimating the illumination map, which is subsequently adjusted maintaining the illumination consistency. Some other techniques [22], [23], [24] use fusion of multiple instances of estimated illumination, thus effectively combining the properties of various approaches.

Besides the traditional methods, modern techniques leverage various deep networks [16] for enhancing low-light images and other similar tasks [10], [11]. Parihar *et al.* [25] studied various deep networks for low-light image enhancement. LLNet [12] presented an autoencoder network that simultaneously performs both contrast enhancement and denoising. GLADNet [13] uses illumination estimation followed by reconstruction of details lost due to illumination adjustment. Wei *et al.* [14] presented a deep network for image decomposition, followed by multi-scale illumination adjustment. It employs dedicated BM3D [26] denoiser to deal with the inherent noise but the denoising operation over-smooths the enhanced image. In contrast to retinex based deep networks, Zero-DCE [27] performs the enhancement by estimating image-specific curve parameters and iteratively applying them on the three channels of the input low-light image. EnlightenGAN [28] is a modern generative adversarial network based approach that works in an unsupervised fashion by training a global-local discriminator network. Nevertheless, a majority of proposed deep networks are computationally expensive, thus paving the way for faster and more practical deep learning based approaches.

III. Methodology

The paper proposes a lightweight retinex theory based deep network, called LRD-Net, for enhancing low-light images. The proposed framework is shown in Fig. 1. The proposed approach mainly includes intrinsic image decomposition into reflectance and illumination followed by adjustments in the estimated components.

The proposed method takes motivation from the Retinex theory [7] which considers an image ($\mathbf{I}$) as made up of two components, namely reflectance and illumination.

$$\mathbf{I} = \mathbf{R} \cdot \mathbf{S} \tag{1}$$

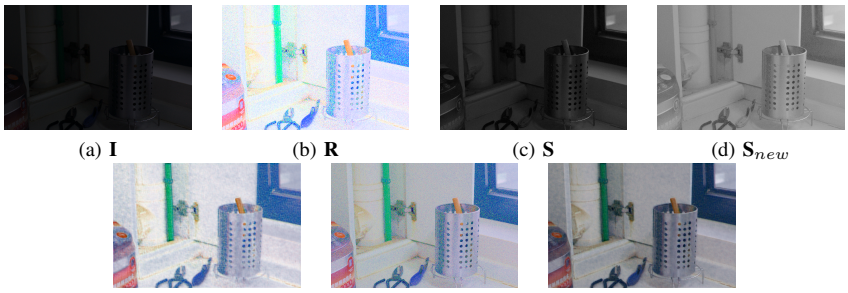
where $\mathbf{R}$ and $\mathbf{S}$ are used to represent reflectance and illumination respectively. The illumination component depends on the lighting conditions while the image is captured. On the other hand, reflectance comprises of the details of the scene and is assumed to be constant for a scene given varying lighting conditions. As illustrated by [21], [22], maximum value of the three color channels contains most of the lighting information and therefore, characterizes the illumination. Thus, the proposed approach employs the max-channel of the input low-light image to estimate the illumination component as follows.

$$\mathbf{S} = \max_{k \in \{R,G,B\}} \mathbf{I}^k \tag{2}$$

Based on the estimated illumination, we compute the reflectance component as:

$$\mathbf{R} = \frac{\mathbf{I}}{\mathbf{S}} \tag{3}$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:30:05 UTC from IEEE Xplore. Restrictions apply.



(a) $\mathbf{I}$ (b) $\mathbf{R}$ (c) $\mathbf{S}$ (d) $\mathbf{S}_{new}$

(e) $\mathbf{R}_{new}$ (f) $\mathbf{R} \cdot \mathbf{S}_{new}$ (g) $\mathbf{R}_{new} \cdot \mathbf{S}_{new}$

Fig. 2: Example to demonstrate the working of proposed method. (a) Input low-light image; (b) Estimated reflectance; (c) Estimated illumination; (d) Gamma corrected illumination; (e) Adjusted reflectance; (f) Initial enhanced image; (g) Final enhanced image.

The reflectance and illumination components for a sample low-light image as estimated by (2) and (3) are shown in Fig. 2. One may observe in Fig. 2(b) that the estimated reflectance comprises of a lot of details that seem lost in the input low-lit image.

The proposed approach subsequently adjusts the illumination using gamma correction [3], which leverages a point-wise operation as shown below.

$$S_{new}(x) = S_{max}\left(\frac{S(x)}{S_{max}}\right)^{\gamma}, \tag{4}$$

where γ is a fixed parameter and S_{max} denotes the maximum intensity value of all the pixels $S(x)$ in the illumination component. The adjusted illumination component for a given image is shown in Fig. 2(d). The recombination of estimated reflectance and gamma adjusted illumination improves the visibility to some extent as displayed in Fig. 2(f). However, the reflectance component loses some structural and textural details in the image decomposition process. Further, gamma correction uses a fixed parameter to adjust the illumination and therefore, results in loss of natural color information. In order to handle such distortions, we present a lightweight convolutional neural network to adjust the reflectance.

The proposed reflectance adjustment network is inspired by [29], and learns a mapping between estimated reflectance ($\mathbf{R}$) and target reflectance ($\hat{\mathbf{R}}$) as obtained by:

$$\hat{\mathbf{R}} = \frac{\hat{\mathbf{I}}}{\mathbf{S}_{new}}, \tag{5}$$

where $\hat{\mathbf{I}}$ is the normal-light ground truth image for input $\mathbf{I}$. The proposed architecture utilizes majorly three operations, including patch extraction, non-linear mapping, and image reconstruction, as shown in Fig. 1. The extracted patches consist of feature maps obtained from the input reflectance. The patches are subsequently mapped to another set of high-quality feature maps, thus removing the distortions introduced in the input reflectance. The reconstruction operation recombines the improved patches to give the final adjusted reflectance component. The proposed LRD-Net employs a three layer architecture performing convolutional and rectified linear unit operations in each of them. Further, we leverage different kernel sizes for the three layers as indicated in Fig. 1.

In order to train the proposed network, we present a multi-objective loss function comprising of three losses, namely MSE loss, content loss, and PSNR loss. The MSE loss calculates the pixel-wise dissimilarity between input and target reflectance estimated as:

$$L_{mse} = \|\mathbf{R}_{new} - \hat{\mathbf{R}}\|_2^2, \tag{6}$$

where $\mathbf{R}_{new}$ is the output of reflectance adjustment network. The content loss [30] is used to preserve the image semantics, and to make sure that the target and adjusted reflectance have similar feature representation.

$$L_{content} = \|\phi(\mathbf{R}_{new}) - \phi(\hat{\mathbf{R}})\|_2^2, \tag{7}$$

where $\phi()$ accounts for the feature map produced after the 4-th Conv layer of VGG-19 network. In order to deal with the inherent noise in the reflectance component, we propose the use of PSNR as a contributing loss. PSNR is widely used for image quality estimation and, therefore, we leverage it to represent the noise level in the output reflectance compared to ground-truth image.

$$L_{PSNR} = -10 \cdot \log_{10}\left(\frac{MAX^2_{\mathbf{R}_{new}}}{MSE}\right), \tag{8}$$

where negative sign is used to reverse the behavior of PSNR, as a greater PSNR value reflects better image quality. The final loss function can be specified as:

$$L_{total} = w_{mse}L_{mse} + w_{content}L_{content} + w_{PSNR}L_{PSNR}, \tag{9}$$

where w_{mse}, $w_{content}$, and W_{PSNR} indicate the balancing weights and are taken as 0.55, 0.3, and 0.15 respectively. Fig. 2 shows the estimated reflectance ($\mathbf{R}$) for a low-light image and the corresponding new reflectance ($\mathbf{R}_{new}$) adjusted by the proposed LRD-Net. Finally, the proposed method recombines the adjusted reflectance ($\mathbf{R}_{new}$) and enhanced illumination ($\mathbf{S}_{new}$) to get the final enhanced image as displayed in Fig. 2(g). It may be noticed that the final enhanced output looks more natural and preserves color information as compared to initial gamma correction based enhanced image.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:30:05 UTC from IEEE Xplore. Restrictions apply.



(a) Input (b) LIME (c) GLADNet (d) Retinex-Net (e) Zero-DCE (f) LRD-Net

Fig. 3: Visual comparison among various methods.



(a) Input (b) LIME (c) GLADNet (d) Retinex-Net (e) Zero-DCE (f) LRD-Net

Fig. 4: Visual comparison among various methods.



(a) Input (b) LIME (c) GLADNet (d) Retinex-Net (e) Zero-DCE (f) LRD-Net

Fig. 5: Visual comparison among various methods.

IV. Experimental Results and Analysis

We use the publicly available LOL dataset [14] that has 500 low/normal light image pairs. We utilize 485 image pairs for training the proposed LRD-Net, and use the remaining 15 pairs along with other datasets, including ExDark [31] and NPE [20] for evaluation. While training, the patch size is taken as 128x128 and batch size is kept as 15. We present the performance comparison of the proposed LRD-Net against various state-of-the-art techniques, including LIME [21], GLADNet [13], Retinex-Net [14], and Zero-DCE [27].

TABLE I: Quantitative comparison using PSNR, SSIM, and NIQE metrics

Method	**Metrics**		
	PSNR	SSIM	NIQE
LIME [21]	13.170	0.720	6.113
GLADNet [13]	15.995	0.757	7.256
Retinex-Net [14]	15.533	0.795	6.452
Zero-DCE [27]	14.395	0.694	8.254
LRD-Net	16.022	0.780	5.970

A. Quantitative Assessment

To assess the performance of the proposed LRD-Net quantitatively, we use three image quality assessment metrics, including PSNR, SSIM, and NIQE [32]. PSNR and SSIM require ground-truth images and thus calculated using images from LOL dataset. On the other hand, NIQE is a blind image quality metric calculated by taking images from all the mentioned datasets. A higher value is desirable for PSNR and SSIM scores, while a good quality image has a low NIQE value. Table 2 gives the values of the three metrics for LIME, GLADNet, Retinex-Net, Zero-DCE, and proposed LRD-Net.

From the values, we can observe that the proposed LRD-Net performs better than other approaches considering the PSNR and NIQE metrics. LIME works worst in terms of PSNR value, while Zero-DCE fails in consideration of NIQE score. The SSIM value is highest for Retinex-Net, and our method falls behind by a margin of 0.015. Thus, our approach shows effective performance in terms of quantitative metrics.

B. Visual Assessment

Quantitative metrics are insufficient in characterizing the visual quality of images as perceived by humans. Thus, we present the visual comparison on some low-light images as enhanced by various methods. Fig. 3 gives the visual comparison on a landscape scene captured in daylight. While all the methods give competitive performance in terms of detail restoration, our method proves its superiority by producing a bright image with better details and contrast. GLADNet and Zero-DCE give fade outputs with poor colors. LIME and Retinex-Net fall behind in terms of brightness improvement as compared to our method.

Fig. 4 shows another example of an indoor image enhanced by various methods. We can see that LIME and Zero-DCE produce darker outputs with lesser details. Retinex-Net gives better performance in terms of restoring the details as compared to LIME and Zero-DCE but loses the naturalness and

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:30:05 UTC from IEEE Xplore. Restrictions apply.

colors as evident in the color palette in Fig. 4(d). GLADNet fails to maintain the color constancy as visible in Fig. 4(c). Our method performs competitively compared to Retinex-Net with regard to details and maintains the naturalness of colors as visible in Fig. 4(f).

Fig. 5 shows an example of a night scene with lightning in the sky. GLADNet and Retinex-Net produce fade results while Zero-DCE gives comparatively darker output. It may be noticed in Fig. 5(b) that LIME over-enhances the bright regions. On the other hand, the proposed LRD-Net sufficiently restores the details without over-enhancement. Therefore, our method is capable of producing visually pleasant outputs.

V. Conclusion

In this paper, we presented a novel lightweight deep learning based method for low-light image enhancement. The proposed method estimates the illumination using max-RGB color channel and computes the corresponding reflectance by element-wise division of input image and estimated illumination. The estimated reflectance is further refined using the proposed lightweight three-layer network, called LRD-Net. The network learns a mapping between estimated and target reflectance in an end-to-end fashion, which is further recombined with gamma adjusted illumination to get the enhanced image. From the experiments, it is evident that the proposed technique demonstrates superior performance to other state-of-the-art approaches, and the lightweight structure gives it an advantage over other methods considering the practical usage.

References

[1] Y.-T. Kim, "Contrast enhancement using brightness preserving bi-histogram equalization," *IEEE transactions on Consumer Electronics*, vol. 43, no. 1, pp. 1–8, 1997.

[2] A. S. Parihar, O. P. Verma, and C. Khanna, "Fuzzy-contextual contrast enhancement," *IEEE Transactions on Image Processing*, vol. 26, no. 4, pp. 1810–1819, 2017.

[3] H. Farid, "Blind inverse gamma correction," *IEEE transactions on image processing*, vol. 10, no. 10, pp. 1428–1433, 2001.

[4] S.-C. Huang, F.-C. Cheng, and Y.-S. Chiu, "Efficient contrast enhancement using adaptive gamma correction with weighting distribution," *IEEE transactions on image processing*, vol. 22, no. 3, pp. 1032–1041, 2012.

[5] A. S. Parihar, "Fuzzy adaptive gamma correction for contrast enhancement," in *2017 International Conference on Intelligent Sustainable Systems (ICISS)*, pp. 625–629, IEEE, 2017.

[6] A. M. Reza, "Realization of the contrast limited adaptive histogram equalization (clahe) for real-time image enhancement," *Journal of VLSI signal processing systems for signal, image and video technology*, vol. 38, no. 1, pp. 35–44, 2004.

[7] E. H. Land, "The retinex theory of color vision," *Scientific american*, vol. 237, no. 6, pp. 108–129, 1977.

[8] D. J. Jobson, Z.-u. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex," *IEEE transactions on image processing*, vol. 6, no. 3, pp. 451–462, 1997.

[9] D. J. Jobson, Z.-u. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on Image processing*, vol. 6, no. 7, pp. 965–976, 1997.

[10] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, and Z. Wang, "Benchmarking single-image dehazing and beyond," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 492–505, 2018.

[11] S. Minaee, Y. Y. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

[12] K. G. Lore, A. Akintayo, and S. Sarkar, "Llnet: A deep autoencoder approach to natural low-light image enhancement," *Pattern Recognition*, vol. 61, pp. 650–662, 2017.

[13] W. Wang, C. Wei, W. Yang, and J. Liu, "Gladnet: Low-light enhancement network with global awareness," in *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, pp. 751–755, IEEE, 2018.

[14] C. Wei, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for low-light enhancement," *arXiv preprint arXiv:1808.04560*, 2018.

[15] L. Shen, Z. Yue, F. Feng, Q. Chen, S. Liu, and J. Ma, "Msr-net: Low-light image enhancement using deep convolutional network," *arXiv preprint arXiv:1711.02488*, 2017.

[16] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[17] A. S. Parihar and K. Singh, "A study on retinex based method for image enhancement," in *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, pp. 619–624, 2018.

[18] K. Singh and A. S. Parihar, "A comparative analysis of illumination estimation based image enhancement techniques," in *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*, pp. 1–5, 2020.

[19] X. Fu, D. Zeng, Y. Huang, X.-P. Zhang, and X. Ding, "A weighted variational model for simultaneous reflectance and illumination estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2782–2790, 2016.

[20] S. Wang, J. Zheng, H.-M. Hu, and B. Li, "Naturalness preserved enhancement algorithm for non-uniform illumination images," *IEEE Transactions on Image Processing*, vol. 22, no. 9, pp. 3538–3548, 2013.

[21] X. Guo, Y. Li, and H. Ling, "Lime: Low-light image enhancement via illumination map estimation," *IEEE Transactions on image processing*, vol. 26, no. 2, pp. 982–993, 2016.

[22] X. Fu, D. Zeng, Y. Huang, Y. Liao, X. Ding, and J. Paisley, "A fusion-based enhancing method for weakly illuminated images," *Signal Processing*, vol. 129, pp. 82–96, 2016.

[23] A. S. Parihar, K. Singh, H. Rohilla, G. Asnani, and H. Kour, "A comprehensive analysis of fusion-based image enhancement techniques," in *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 823–828, IEEE, 2020.

[24] H. Rohilla, G. Asnani, K. Singh, and A. S. Parihar, "Low-light image enhancement using multi-exposure sequence generation and image fusion," *International Journal of Advanced Science and Technology*, vol. 29, pp. 4481 – 4490, May 2020.

[25] A. S. Parihar, S. Singhal, S. Nanduri, and Y. Raghav, "A comparative analysis of deep learning based approaches for low-light image enhancement," in *2020 5th IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*, pp. 1–5, IEEE, 2020.

[26] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising with block-matching and 3d filtering," in *Image Processing: Algorithms and Systems, Neural Networks, and Machine Learning*, vol. 6064, p. 606414, International Society for Optics and Photonics, 2006.

[27] C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, and R. Cong, "Zero-reference deep curve estimation for low-light image enhancement," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1780–1789, 2020.

[28] Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, and Z. Wang, "Enlightengan: Deep light enhancement without paired supervision," *IEEE Transactions on Image Processing*, vol. 30, pp. 2340–2349, 2021.

[29] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 2, pp. 295–307, 2015.

[30] A. Ignatov, N. Kobyshev, R. Timofte, K. Vanhoey, and L. Van Gool, "Dslr-quality photos on mobile devices with deep convolutional networks," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3277–3285, 2017.

[31] Y. P. Loh and C. S. Chan, "Getting to know low-light images with the exclusively dark dataset," *Computer Vision and Image Understanding*, vol. 178, pp. 30–42, 2019.

[32] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a "completely blind" image quality analyzer," *IEEE Signal processing letters*, vol. 20, no. 3, pp. 209–212, 2012.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:30:05 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Fifth International Conference on Trends in Electronics and Informatics (ICOEI).
IEEE Xplore Part Number:CFP21J32-ART; ISBN:978-1-6654-1571-2

Machine Learning and IoT based solutions for detection of arrhythmia using ECG signals

Vinod Kumar
Dept. of CSE
Delhi Technological University
New Delhi, India
vinod_k@dtu.ac.in

Awishkar Ghimire
Dept. of CSE
Delhi Technological University
New Delhi, India
awishkar.ghimire@gmail.com

Hitesh Kumar Hoon
Dept. of CSE
Delhi Technological University
New Delhi, India
hitesh.hoon101@gmail.com

2021 5th International Conference on Trends in Electronics and Informatics (ICOEI) | 978-1-6654-1571-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICOEI51242.2021.9452907

***Abstract*—The World Health Organization recognizes heart diseases as being the leading cause of death globally bringing about 17.9 million deaths annually. Arrhythmia is one of the most common heart diseases and hence it is of quintessential importance that there be computational systems that are able to automatically diagnose arrhythmia. ECG is a way of detecting the heart beat using electrical changes in the skin caused by the cardiac cycle. ECG can be used to automatically diagnose various heart diseases including arrhythmia. It is very well known that artificial intelligence is known as the 4th industrial revolution and various subfields of artificial intelligence such as machine learning and deep learning have been used to solve problems across a wide range of fields. Also, IoT is a technology that is increasingly being used to solve various real world problems computationally. In this paper we have reviewed the most impactful and noteworthy research papers in recent years that deal with the subject matter of arrhythmia detection using ECG signals based on machine learning and IoT.**

***Index Terms*—ECG, Arrhythmia, Machine Learning, Deep Learning, IoT, Long Short Term Memory(LSTM), Support Vector Machine(SVM), MIT-BIH Arrhythmia database.**

I. Introduction

Arrhythmia can be defined as a problem regarding the rate of rhythm of the heart beat of a person. It means that the heart is beating at a faster rate or a slower rate than normal. It is also characterized by an irregular heartbeat. When the heart beats too fast the term given to that particular type of arrhythmia is tachycardia whereas when the heart beats faster than average the term given to it is bradycardia. One of the most important types of arrhythmia is atrial fibrillation. Atrial fibrillation is a type of quivering, rapid heartbeat that can lead to heart attacks, strokes, blood clots and heart palpitation. Many factors can cause arrhythmia such as heart attack, coronary artery disease, smoking, stress, congenital heart defects among other causes. Treatment to restore the normal heart rhythm of an arrhythmia patient includes medicines, an implantable cardioverter-defibrillator as well as surgery in the most extreme cases. Symptoms of arrhythmia can include heart palpitations, lightheadedness, passing out, shortness of breath. It can cause heart failures, strokes and death in the most extreme cases. Arrhythmia can be broadly classified into 4 main classes namely, bradyarrhythmias, supraventricular tachycardias, ventricular arrhythmias, and extra beats[14].

According to the World Health Organization, heart diseases also known as cardiovascular diseases is the leading cause of death globally, bringing about 17.9 million deaths per year. Cardiovascular diseases are a group of disorders of the heart and various blood vessels around the heart and include diseases like coronary artery disorder, cerebrovascular disease. Among all the deaths caused due to cardiovascular diseases four out of five such deaths are due to heart attacks and strokes, and about one third of the deaths happen prematurely to people below the age of 70 [3]. All cardiovascular disease can cause different types of arrhythmia. Specifically coronary artery disease can cause a certain type of a very dangerous arrhythmia called the atrial fibrillation. Thus it is of paramount importance to create systems specifically computational systems that can accurately diagnose arrhythmia in patients.

The most common and effective way of diagnosing arrhythmia is through the use of ECG. Electrocardiogram (ECG) is a graph in which the y axis is the voltage and x axis is the time. It is a voltage vs time graph. The voltage recorded in the ECG is driven by the electrical activity of the heart. The voltage is recorded using electrodes placed on the skin. These electrodes can detect the electrical changes in the skin that is caused by the cardiac cycle(heartbeat). The voltage change is caused by the depolarization followed by repolarization of the cardiac muscle. There are different types of ECG. The most common one is the 12 lead ECG which consists of 10 different electrodes placed on the skin of the patient. The electrodes are placed in different parts of the body which includes the limbs and the chest. The different electrodes detect the electrical activity of the heart in different angles giving different perspectives of the same cardiac cycle. A cardiac cycle in an ECG consists of three main components, namely the P wave, the QRS complex and the T wave. The P wave illustrates the depolarization of the atria, the QRS complex illustrates the depolarization of the ventricles and the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:54:27 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Fifth International Conference on Trends in Electronics and Informatics (ICOEI).
IEEE Xplore Part Number:CFP21J32-ART; ISBN:978-1-6654-1571-2

T wave illustrates the ventricular repolarization . Usually for basic heart monitoring 1 lead ECG is used which consists of only two electrodes and it detects the electrical activity of the heart from only a single angle of view. Lead 1 ECG has been gaining popularity as it can be used easily and is cheaper. It is being extensively used for basic heart monitoring .

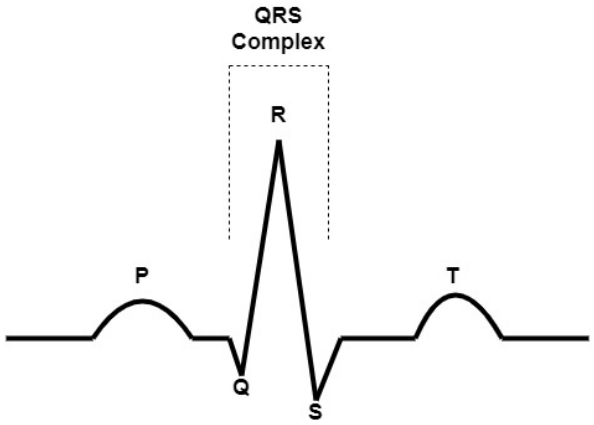


Fig. 1. Representation of the ecg signal of a regular heartbeat

Many research has been done in the field of automated diagnosis of different arrhythmias using machine learning and artificial intelligence on ECG data. Machine learning is a subfield of artificial intelligence and artificial intelligence has been successfully used to solve a wide array of tasks. Machine learning has been and is being used extensively in various businesses particularly to improve the growth of a business[6]. Machine learning and artificial intelligence is also being successfully used in the healthcare industry to combat various different diseases such as alzheimer's, parkinson's and also COVID-19 [22][21][7]. Artificial intelligence and machine learning is also being used successfully in various disciplines of engineering such as civil engineering, electrical engineering and such. It is being used to crack detection in roads and also in niche fields such as urban planning [13][12]. It is also being used for the detection of illegal hacking [8]. The current world is heavily reliant on the field of artificial intelligence and machine learning and plethora of research has been conducted on various fields using this technology. So it is of no surprise that this technology can also be used in the diagnosis of heart disease. It is extensively being used for diagnosing arrhythmia using ECG signals and this paper attempts to review some of the most notable and worthy papers in this particular subject.

Internet of things(IoT) is another technology that is sweeping the current world. It finds applications in various different areas such as healthcare, business, communication systems and other fields[5]. IoT helps in making environment smarter and improve living standards, for example, with the help of IoT people can make their home smart, generate smart energy, build smart cities[20]. The internet of things is a technology that can be broadly defined as a network of intelligent devices working together in harmony to solve a particular problem. The architecture mainly consists of sensors, computational processors, microcontrollers communicating over a network. Many IoT architectures also incorporate cloud based strategy. IoT has also been extensively used in the detection of arrhythmia using ECG signals. Recent research work are performed in IoT based monitoring systems for health care which proposes management of big data using IoT environment helping in finding big data based solutions for various problems in health care[17]. It is mainly incorporated in remote monitoring of patients and is used for real time ECG analysis. In this paper we attempt to review the most recent advancements in the use cases of IoT for arrhythmia detection.

II. Related Works

One of the most impactful papers on ECG arrhythmia detection has been by Rajpurkar et al. Rajpurkar et al [18] have developed a machine learning algorithm that has been shown to work better than board certified cardiologist in detecting a wide range of different arrhythmias using an ECG recorded from a single lead ECG machine. They have used a convolution neural network model consisting of 33 layers of convolution after which comes a fully connected layer and a softmax layer, so in total their model has 34 layers. The input consists of a sequence of numbers denoting the ECG signal and an output consists of arrhythmia classes of the different heartbeats in the input. The model makes use of batch normalization and the activation function used in the model is the Relu function. The model also makes use of a dropout algorithm to reduce overfitting. The data set Rajpurkar et al have used was collected by the team and it consists of 64,121 ECG records which were detected from 29, 163 patients., and were sampled at 200Hz. The records were annotated by a group of experts. The number of output classes in the study were 14, out of which 12 were different arrhythmia classes, one was the sinus rhythm which is the normal heartbeat and one was noise. The data was partitioned into the training set and test set and they tested their model against board certified cardiologist and it was found that their model performed better. The precision of their model turned out to be 0.8 whereas the precision of the board certified cardiologist came out to be 0.723 . This is a very relevant study because it shows that computational models are better at detecting heart arrhythmias using single lead ECG than human experts.

Yao et al[25] have proposed an attention based time incremental convolutional neural network(ATI-CNN) for the detection of arrhythmia using 12 lead ECG. They have used the ATI-CNN model to address the issues of convolutional neural networks(CNN) in detecting arrhythmia such as CNN's inability to handle varied length ECG signals. Their proposed model consists of a 13 layer CNN module , a 2 layer LSTM module and an attention module one after another and hence the model achieves both spatial and temporal fusion of information from ECG signals. The dataset that was used in this study was from the 1st China Physiological Signal challenge and it consisted of 12 lead ECG recordings ranging from 6 seconds to 60 second. There were 6877 ECG recordings in

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:54:27 UTC from IEEE Xplore. Restrictions apply.

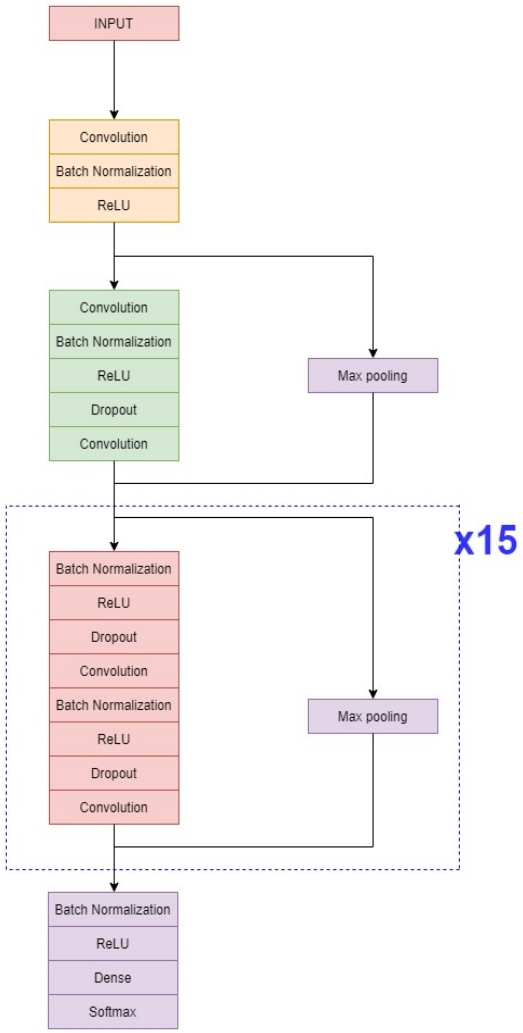


Fig. 2. The model proposed by Rajpurkar et al [18] for the detection of arrhythmia from one lead ECG.

the training set and 2954 ECG recordings in the test set. The data was sampled at 500Hz. The number of output classes that were considered in this study was 9 including, normal beat. The precision of their model turned out to be 0.826 and it's accuracy was 81.2% which was 7.7% more than a standard 16 layer CNN named VCGNet.

Yildirim et al[26] have proposed an efficient and real-time classification model for detecting 17 unique classes of arrhythmia using 1-lead ECG(MLII). they used 10-second long fragments of ECG signal for classification. They used a 16-layer deep 1D-Convolution Neural Network along with BatchNormalization and Max Pooling. Input passes through 1D convolution layer and then the outputs are normalized with the help of Batch Normalization layer, then normalized output passes through 1D max pooling layer. These operations are repeated until the 14th layer. The 14th layer of this deep

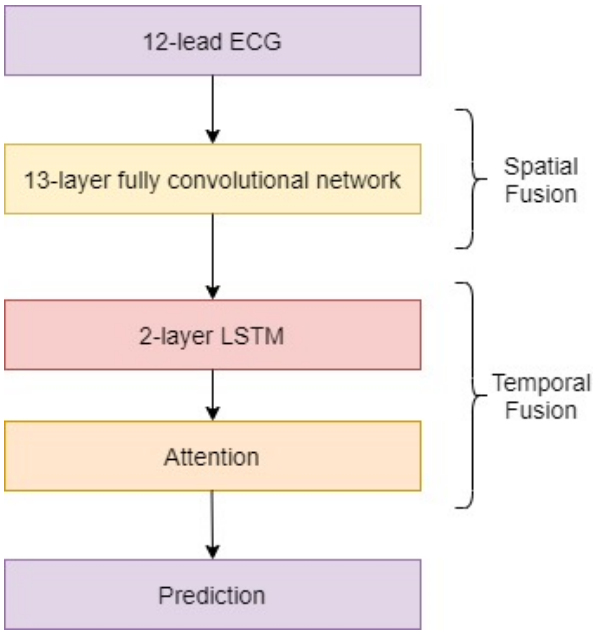


Fig. 3. The model proposed by Yao et al[25] for arrhythmia detection using 12-lead ECG

neural network is a flattened layer which transforms multi-dimensional input feature vector into 1D output data. The feature vectors are then passed to a Dense layer of size 512 units. The final layer uses Softmax as the activation function and predicts which arrhythmia class the input belongs to. The dataset used is the MIT-BIH Arrhythmia dataset. Among the 1000 ECG signal snippets, the training set includes 700 samples and the validation and testing set contains 150 samples each. After training for 50 epochs, a training accuracy of 100% and a validation accuracy of 86.67% were obtained. The proposed 16 layered deep 1D-CNN results in an accuracy of 91.33% on testing data and the classification time per single sample was 0.015 second.

Wang et al[23] have proposed a unique model for the purpose of arrhythmia detection. They have used deep multi scale fusion neural networks that can effectively classify multiple arrhythmias. The model they have used in their research consists of three key components. The first is a backbone network that is able to learn shared low level features. The second is a collection of sub-networks that can effectively learn scale specific high level signal features and different scales convolution kernel collaboratively. The third component of the model is a multi scale features fusion that is used for integrating all of the different features from the sub networks in the previous component and further uncovering correlated information from different scales using attention. Two different data sets have been used in this particular research. The first data set is the CPSC_2018 data set which consists of 6877 12 lead ECG records for training and 2954 ECG records for testing, sampled at 500 Hz and ranging from 6 to 60 seconds. This data set includes 9 different types of heartbeat

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:54:27 UTC from IEEE Xplore. Restrictions apply.

namely, normal beat and 8 different types of arrhythmia. The second data set that has been made use of in this paper is the PhysioNet/CinC_2017 data set. This data set consisted of 8, 528 records of single lead ECG data sampled at 300 Hz. The lengths of the recordings ranges between 2700-18, 300 data points. The output of these ECG signals were labeled with four classes namely, sinus rhythm, atrial fibrillation, alternative rhythm and noisy recordings. Their model achieved an F1 score and precision of 0.828 and 0.838 respectively on the CPSC_2018 data set. On the CINC_2017 data set their model achieved a F1 score of 0.841 and a precision of 0.856.

Miquel et al[1] developed a fast and fully automatic machine learning ECG arrhythmia classifier, that only requires a single ECG lead. They used ensemble learning to their advantage helping in training the classifier with less time. Their ECG arrhythmia classifier is evaluated on MIT-BIH AR data set and (MIT-BIH) AHA data set. This classifier is an ensemble of Echo State Networks, which uses raw ECG waveforms and heartbeats time interval for input. ESNs have recurrent connections which allow them to take time dependencies between two heartbeats into account.The MIT-BIH AR(arrhythmia) database consists of 48 half-hour ECG records sampled at 360 Hz. The AHA database contains 3 hours long, 154 ECG records. Each data set is divided into two sets, for training and for testing. The data is resampled at 250 Hz with a bandwidth of 0.5Hz-35Hz segmented around annotated position with a size of 240ms. The classifier used in this paper is based upon Echo State Network(ESN), which is considered as implementation of Reservoir Computing(RC). RC is a machine learning paradigm which consists of input, reservoir and output layer. In ESN the reservoir is a Recurrent Neural Network which allows context-dependent computations.The reservoir consists of sigmoid function which shows optimal results for classification. The output is obtained from the ensemble by majority voting over individual ESN output. The classifier used in this paper is an ensemble of thirty ESNs following ring structure. The proposed classifier classifies the input into two classes. For MIT-BIH AR data set the accuracy of the classifier is 98.6% and for MIT-BIH AHA data set the accuracy achieved by the classifier is 97.8%.

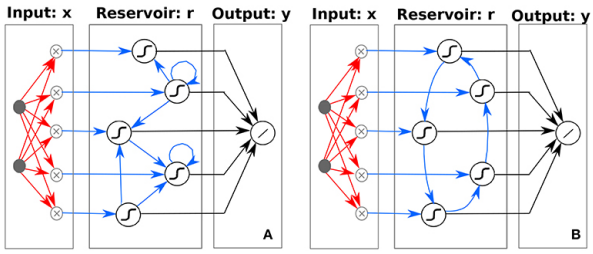


Fig. 4. The Reservoir Computing method proposed by Miquel et al[1] for arrhythmia detection using ECG signals.

Sannino et al[19] proposed a Deep Neural Network of 7 hidden layers for heartbeat classification. The data set used in this paper is MIT-BIH Arrhythmia data set. This data set is used to create a new preprocessed data set. Noises are eliminated, position of peaks are detected, signals are segmented into single beats classified as normal or abnormal. After these preprocessing steps the new data set consists of a total of 84,615 items, out of these 66,750 are normal beats(N), 2288 are abnormal beats and rest 14,828 represents unclassifiable beats(Q). These unclassifiable beats are dropped and only 2288 randomly selected normal beats out of 66,750 normal beats are considered so that the classification model does not suffer from bias. Thus the final data set consists of 4576 items. This data is used to create a Deep Neural Network which predicts two classes: normal beat and abnormal beat. After various manual configurations, the DNN with best classification performance contains seven hidden layers, with 5, 10, 30, 50, 30, 10 and 5 neurons respectively. The activation function for the hidden layer is ReLU and for the output layer it is Softmax. The proposed classifier has an accuracy of 100% on training data and an accuracy of 99.09% on test data, while the sensitivity and specificity for training data is 100% and for testing data it is 98.55% and 99.52% respectively.

Isin et al[11]have proposed a transfer learning method for detection of arrhythmia from ECG signals. A very popular convolutional neural network named AlexNet is used for feature extraction and then the extracted features is fed to a simple feed forward neural network that is trained using backpropagation. The workflow in their study undergoes various steps. The first is data acquisition in which they use a very popular dataset known as the MIT-BiH dataset. Then various signal smoothing algorithms are applied to the signal. Further Pan-Tompkins algorithm is used for QRS detection after which the signal is fed onto AlexNet for feature extraction. They test 3 different neural networks for classification and the highest accuracy that was obtained was 92.4%.

Petmezas et al[16] proposed a hybrid CNN-LSTM approach for arrhythmia detection. Atrial fibrillation is an arrhythmia that can cause failure of heart. There are four classes of ECG beats predicted by the proposed architecture. The dataset used is MIT-BIH Atrial Fibrillation database, which includes 23 long term 2-lead ECG readings sampled at 250 Hz. The data preprocessing step consists of removing unwanted noises, beat segmentation. The preprocessed data is then fed to Convolutional Neural Networks. The CNN is used to extract meaningful ECG features which are then inputted to a Long Short-Term Memory(LSTM) model, this results in memorization of temporal dynamics which help in more accurate classification of ECG beats. To deal with the imbalance data focal loss is implemented. The proposed hybrid CNN-LSTM architecture has thirteen hidden layers other than input and output layers. First, ECG signal is passed through CNN parts which have triplets of layers with each triplet including Convolutional, Batch Normalization and MaxPool layer. The CNN part is responsible for feature extraction as well as dimensionality reduction. The results from CNN part move to LSTM part. The LSTM part is responsible for memorizing time related dependencies between data. The LSTM part includes LSTM layer, flatten, fully-connected, dropout and output layers. Activation function for the convolutional and fully connected layer

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:54:27 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Fifth International Conference on Trends in Electronics and Informatics (ICOEI).
IEEE Xplore Part Number:CFP21J32-ART; ISBN:978-1-6654-1571-2

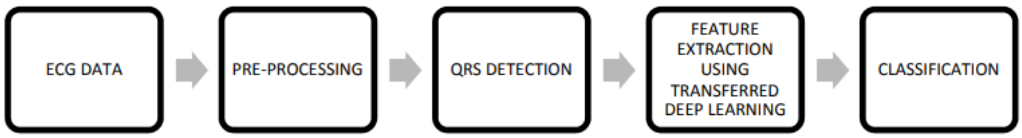


Fig. 5. The transfer learning method proposed by Isin et al[11] for arrhythmia detection using ECG signals.

is ReLU and for the output layer it is Softmax. Then a 10-fold CV is used for model evaluation. The training process ends in 100 epochs resulting in an overall sensitivity of 97.87% and an overall specificity of 99.29%.

Hong et al[10] have performed a study on arrhythmia detection using single lead ECG signals. In their study they have attempted to combine hand engineered features and features extracted from deep neural networks. The combined features are then further inputted to a classifier for classification of the detected arrhythmia. They have used various types of engineered features such as statistical features, signal features, morphological features. They have used various different types of deep neural networks like CNN, RNN, ResNet for feature extraction. The extracted and the engineered features were then trained and tested using different classifiers such as softmax, decision tree, XGBoost. Ensemble XGBoost. The best performance was given by ensemble XGBoost. The performance measure used by them in their study was the F-score and the F-score obtained by their model was 0.7605. The data set used in this study was obtained from PhysioNet/Computing in Cardiology Challenge 2017. This data set consisted of 8, 528 records odf single lead ECG data sampled at 300 Hz. The lengths of the recordings ranges between 2700-18,300 data points. The output of these ECG signals were labeled with four classes namely, sinus rhythm, atrial fibrillation, alternative rhythm and noisy recordings.

Warrick et al[24] have proposed a model which combines CNN architecture and LSTM architecture for the detection of arrhythmia using ECG signals. The model consists of 2 main components namely the representative learning component and the sequence learning component. The representative learning component consists of 1D CNN layer which is used for feature extraction, and the output of this layer is fed onto the sequence learning layer which consists of a stack of three LSTM layers. After the LSTM layer the data is fed onto a softmax layer for classification. The data set they used was from the PhysioNet Challenge 2017 which is a fairly common and standard data set. The number of output classes used in this study was 4. The primary performance measure they used to evaluate their model. They conducted a series of experiments and the highest F1 score that was obtained was 0.83.

III. Further Discussion

Moghadas et al[15] have proposed an IoT based framework for the automatic detection of arrhythmia using remote monitoring of a patient's health. Their study makes use of fog computing The patient monitoring system is based on fog computing. Fog computing is an IoT architecture in which the edge devices are responsible for a significant amount of computation and storage. In their Iot based systems three main hardware components are used. The first is the Arduino Uno which is a microcontroller, The second is the Raspberry Pi 3 B+. which is a computer that is responsible for computation of the pre trained machine learning algorithm. The third is the AD8232 board which is responsible for detecting ECG signals and heartbeat of the patient. The machine learning algorithm they used for the classification was the KNN algorithm. The data set they used to train the machine learning algorithm was the UCI data set for arrhythmia. In the particular dataset 16 different output classes were present and the model used in the study gave an accuracy of 94.44%. In their particular IoT model, the sensor sensed the ECG data and the computing was done by the processors. The data was then sent to the cloud for storage. In case of an emergency or heart attack the patient's family or physician was notified through an alert system.

Jinyuan et al[9] have proposed a method for detection of arrhythmia from IoT-based ECGs. The proposed method consists of two parts, first is data cleaning part and the other one is heartbeat classification part.The data set used is MIT-BIH database. The proposed model predicts five classes. Input data first passes through a data cleaning module in which ECG recordings are re-sampled to 360 Hz, then these signals are filtered of any unwanted noise and then 230 heartbeat samples are obtained using heartbeat segmentation. This paper proposes two solutions: Dynamic Heartbeat Classification with Adjusted Features(DHCAF) and Multi-channel Heartbeat Convolution Neural Network(MCHCNN). DHCAF method is based on feature selection. DHCAF contains 4 stages: first being Feature extraction , the second step is building a classifier training pool with various classifiers(eg.,SVM, decision tree, etc), in the third step most competent classifiers are selected for making predictions and then at the last step the aggregated results from previous steps are refined. MCHCNN takes in two inputs, one being an ECG beat and other being RR-intervals. MCHCNN performs parallel convolutions using 3 channels on the input beats to retrieve frequency and time related information of ECG heartbeats. After every convolution operation there is a batch normalization step and the activation function used is ReLU. The features learned from three parallel implemented convolution layers are integrated and then passed to the pooling layer. Then a Rhythm Integration layer is

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:54:27 UTC from IEEE Xplore. Restrictions apply.

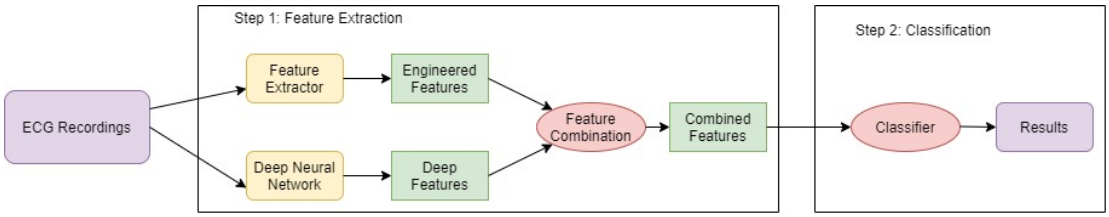


Fig. 6. The model proposed by Hong et al[10] for arrhythmia detection using single lead ECG

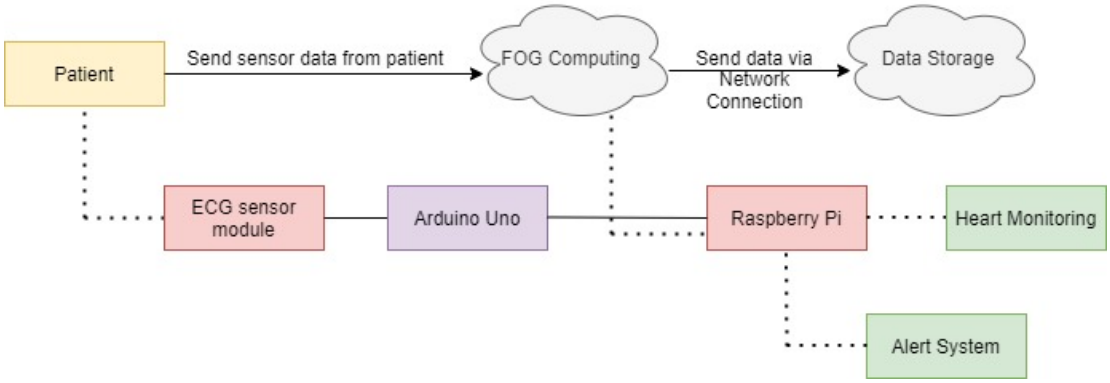


Fig. 7. Iot architecture used in the study done my Moghadas et al[15]

used to join input heart rhythms(RR-intervals) and the learned features. Next, a dense layer followed by a Softmax layer predicts the class to which the beats belong. The proposed DHCAF method has an accuracy of 91.4% and the proposed MCHCNN method has an accuracy of 93%.

Dimitra et al[2] proposed an algorithm for arrhythmia detection by ECG signal analysis and implemented it on an IoT-based platform. The dataset used is MIT-BIH Arrhythmia database. The IoT platform used is Intel's Galileo board, with Linux operating system, 400MHz processor and a RAM of 256MB. Application implemented on Galileo board takes a digitized ECG signal as input and these signals are read at 350 samples/second with a batch size of 3000 samples. First a bandwidth filter is written in C which is used to filter the signals at 1-50 Hz. Next, WFDB function wqrs() is implemented in C which gives back the starting point of QRS complex. Next, heartbeat segmentation is performed on the ECG signal. Next, Discrete Wavelet Transform method is used for feature extraction. After getting feature vectors a binary classifier is implemented which predicts beats as Normal and Abnormal. They used 104581 heartbeats from 45 selected ECG readings for classification. The LIBSVM library of Matlab interface also includes C implementation of svmpredict() which converts SVM model into format used in C implementation. The proposed signal analysis and classification architecture detected 10 heartbeats for every 3000 sample inputs, and in best case total time of execution is less than 9 seconds which shows that it can be used in real-time. The classification accuracy achieved by the proposed algorithm is above 97% while the best case accuracy is 98.9% for 16 feature vectors and 2493 support vectors.

Lakshmi et al[4] proposed an IoT-based cardiac arrhythmia detection system using ECG signals. The data set used for training and validation of the classifier is obtained from Physionet ATM database. ECG signals are collected from various data sets: eighteen ECG readings from MIT-BIH Normal Sinus Rhythm Database(NSRDB), twenty two ECG recordings from MIT-BIH Malignant Ventricular Ectopy database, thirty five recordings from CU Ventricular Tachycardia Database and seventy eight ECG signals from MIT-BIH Supraventricular Arrhythmia Database. To acquire ECG signals they used an Arduino controller which receives data from ECG sensor AD8232. The AD8232 sensor reads electrical signals from the heart and then the Arduino controller gets that data and transmits it to the processing system. To detect QRS complexes, the Pan Tompkins algorithm is used. Feature extraction of raw ECG signal gives statistical features and dynamic features. The model for classification is implemented in MATLAB and the classification app consists of various types of SVM classifiers. The proposed algorithm with only statistical features set has an effectiveness of 88.9% and the proposed algorithm with dynamic features has an effectiveness of 90.8% while a combination of these two features has an effectiveness of 92.2%.

IV. Conclusion and Future Works

In this paper we have reviewed various recent and impactful papers on arrhythmia detection using various artificial intelligence and IoT based techniques. As stated above heart diseases are one of the leading causes of human deaths and

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:54:27 UTC from IEEE Xplore. Restrictions apply.

TABLE I
SUMMARY OF WORKS DONE IN THE FIELD OF ARRHYTHMIA DETECTION

Author	Date of Publication	Technology Used	Dataset Used	Performance Measure
Rajpurkar et al[18]	2017	Convolutional Neural Network	Novel data set collected by the authors	Precision: 0.8, F1 score: 0.776
Yao et al[25]	2020	Attention based time incremental Convolution Neural Network	1st China physiological Signal challenge data set	Precision: 0.826, F1 score: 0.812
Hong et al[10]	2019	DNN + machine learning	PhysioNet/ computing in Cardiology Challenge data set	F1 score: 0.7605
Wang et al[23]	2020	Deep multi scale fusion Neural Network	physioNet/CiNC 2017 and CPSC_2018 data set	F1 score(CPSC): 0.828, F1 score(CINC): 0.841
Isin et al [11]	2017	Transfer Learning using AlexNet + Neural Network	MIT-BiH arrhythmia dataset	Accuracy: 92.4%
Warrick et al[24]	2017	Convolutional Neural Network + Long short Term Memory	PhysioNet Challenge 2017	F1 score: 0.83
Moghadas et al[15]	2019	Internet of Things + K-nearest neighbours	UCI dataset for arrhythmia	Accuracy: 94.44%
Yildirim et al[26]	2018	1D-Convolution Neural Network	MIT-BIH Arrhythmia data set	Accuracy: 91.33%
Miquel et al[1]	2019	Echo State Network based classifier	MIT-BIH AR dataset and MIT-BIH AHA dataset	Accuracy(AR): 98.6%, Accuracy(AHA): 97.8%
Sannino et al[19]	2018	Deep Neural Network	MIT-BIH Arrhythmia database	Sensitivity:98.55%, Specificity:99.52%
Petmezas et al [16]	2021	Hybrid Convolution Neural Network and Long Short Term Memory	MIT-BIH Arrhythmia database	Sensitivity:97.87%, Specificity:99.29%
Dimitra et al [2]	2016	Internet of Things + Support Vector Machines	MIT-BIH Arrhythmia dataset	Accuracy : 97%
Lakshmi et al [4]	2019	Internet of Things + Support Vector Machines	MIT-BIH arrhythmia data set	Accuracy: 92.2%
Jinyuan et al [9]	2020	Multi-channel Heartbeat Convolution Neural Network(MCHCNN) and Dynamic Heartbeat Classification with Adjusted Features(DHCAF)	MIT-BIH Arrhythmia dataset	Accuracy(DHCAF): 91.4%, Accuracy (MCHCNN): 93%

it is quintessential for computational systems to be able to accurately diagnose various heart diseases. With this paper we have attempted to give a glimpse of the major advancements in the field of arrhythmia detection in recent years. We have reviewed various papers on the subject matter and succinctly written about them in a technical manner. This paper will be useful for any researcher or academician venturing into the world of arrhythmia detection using recent computational technologies. The recent landscape of the subject matter mostly consists of deep neural networks such as deep convolutional neural networks, RNN or a hybrid deep neural network comprising various different neural networks. These networks have successfully been shown to give very accuracy in the field of arrhythmia detection. As the amount of ECG data grows and artificial intelligent technologies advance, it is sure that computational systems that are capable of 100% accuracy will be created in the near future. Perhaps one important application future researchers can work on is the early detection of arrhythmia, this is to be able to detect different types of arrhythmia even before it occurs in patients. The ultimate goal in the field of automatic arrhythmia diagnosis using ECG signals would be to create an end to end system that is capable of monitoring the heart beat of a patient all the time and give 100% accurate diagnosis of any little change that is happening in the heart of the patient. In the future artificially intelligent systems will probably make the role of a cardiologist obsolete as the system itself will correctly diagnose and treat a patient.

REFERENCES

[1] Miquel Alfaras, Miguel C Soriano, and Silvia Ortın. “A fast machine learning model for ECG-based heartbeat classification and arrhythmia detection”. In: *Frontiers in Physics* 7 (2019), p. 103.

[2] Dimitra Azariadi et al. “ECG signal analysis and arrhythmia detection on IoT wearable medical devices”. In: *2016 5th International conference on modern circuits and systems technologies (MOCAST)*. IEEE. 2016, pp. 1–4.

[3] *Cardiovascular diseases*. URL: https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:54:27 UTC from IEEE Xplore. Restrictions apply.

[4] R Lakshmi Devi and V Kalaivani. "Machine learning and IoT-based cardiac arrhythmia diagnosis using statistical and dynamic features of ECG". In: *The Journal of Supercomputing* (2019), pp. 1–12.

[5] Elgar Fleisch, Markus Weinberger, and Felix Wortmann. "Business models and the internet of things". In: *Interoperability and open-source solutions for the internet of things*. Springer, 2015, pp. 6–10.

[6] Awishkar Ghimire et al. "Accelerating business growth with big data and artificial intelligence". In: *Proc. 4th Int. Conf. IoT Soc. Mobile, Anal. Cloud, ISMAC 2020*. November. 2020, pp. 441–448. ISBN: 9781728154640. DOI: 10.1109/I-SMAC49090.2020.9243318.

[7] Awishkar Ghimire et al. "AI and IoT Solutions for Tackling COVID-19 Pandemic". In: *Proc. 4th Int. Conf. Electron. Commun. Aerosp. Technol. ICECA 2020*. November. 2020, pp. 1083–1092. ISBN: 9781728163871. DOI: 10.1109/ICECA49313.2020.9297454.

[8] Awishkar Ghimire et al. "Machine Learning Approach Based on Hybrid Features for Detection of Phishing URLs". In: *2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*. IEEE. 2021, pp. 954–959.

[9] Jinyuan He et al. "A framework for cardiac arrhythmia detection from IoT-based ECGs". In: *World Wide Web* 23.5 (2020), pp. 2835–2850.

[10] Shenda Hong et al. "Combining deep neural networks and engineered features for cardiac arrhythmia detection from ECG recordings". In: *Physiological measurement* 40.5 (2019), p. 054009.

[11] Ali Isin and Selen Ozdalili. "ScienceDirect ScienceDirect Cardiac arrhythmia detection using deep learning". In: *Procedia Comput. Sci.* 120 (2018), pp. 268–275. ISSN: 1877-0509. DOI: 10.1016/j.procs.2017.11.238. URL: https://doi.org/10.1016/j.procs.2017.11.238.

[12] Avinash Kumar Jha et al. "A Review of AI for Urban Planning: Towards Building Sustainable Smart Cities". In: *Proc. 6th Int. Conf. Inven. Comput. Technol. ICICT 2021*. January. 2021, pp. 937–944. ISBN: 9781728185019. DOI: 10.1109/ICICT50816.2021.9358548.

[13] Arunish Kumar et al. "Crack Detection of structures using Deep learning frameworks ." In: 13 (2020), pp. 1–24. DOI: 10.1109/ICISS49785.2020.9315949.

[14] Shanthi Mendis et al. *Global atlas on cardiovascular disease prevention and control*. World Health Organization, 2011.

[15] Ehsan Moghadas, Javad Rezazadeh, and Reza Farahbakhsh. "An IoT Patient Monitoring based on Fog computing and Data Mining : Cardiac Arrhythmia Usecase". In: *Internet of Things* (2020), p. 100251. ISSN: 2542-6605. DOI: 10.1016/j.iot.2020.100251. URL: https://doi.org/10.1016/j.iot.2020.100251.

[16] Georgios Petmezas et al. "Automated Atrial Fibrillation Detection using a Hybrid CNN-LSTM Network on Imbalanced ECG Datasets". In: *Biomedical Signal Processing and Control* 63 (2021), p. 102194.

[17] Jennifer S Raj. "A Novel Information Processing in IoT Based Real Time Health Care Monitoring System". In: *Journal of Electronics* 2.03 (2020), pp. 188–196.

[18] Pranav Rajpurkar et al. *Cardiologist-level arrhythmia detection with convolutional neural networks*. 2017. arXiv: 1707.01836.

[19] Giovanna Sannino and Giuseppe De Pietro. "A deep learning approach for ECG-based heartbeat classification for arrhythmia detection". In: *Future Generation Computer Systems* 86 (2018), pp. 446–455.

[20] S Smys. "A Survey on Internet of Things (IoT) based Smart Systems". In: *Journal of ISMAC* 2.04 (2020), pp. 181–189.

[21] Surendrabikram Thapa et al. "Data-Driven Approach based on Feature Selection Technique for Early Diagnosis of Alzheimer's Disease". In: *Proc. Int. Jt. Conf. Neural Networks*. 2020. ISBN: 9781728169262. DOI: 10.1109/IJCNN48605.2020.9207359.

[22] Surendrabikram Thapa et al. "Feature Selection Based Twin-Support Vector Machine for the diagnosis of Parkinson's Disease". In: *2020 IEEE 8th R10 Humanitarian Technology Conference (R10-HTC)*. 2020.

[23] Ruxin Wang, Jianping Fan, and Ye Li. "Deep Multi-Scale Fusion Neural Network for Multi-Class Arrhythmia Detection". In: *IEEE J. Biomed. Heal. Informatics* 24.9 (2020), pp. 2461–2472. ISSN: 21682208. DOI: 10.1109/JBHI.2020.2981526.

[24] Philip Warrick and Masun Nabhan Homsi. "Cardiac arrhythmia detection from ECG combining convolutional and long short-term memory networks". In: *Comput. Cardiol. (2010)*. Vol. 44. 2017, pp. 1–4. DOI: 10.22489/CinC.2017.161-460. arXiv: 1801.10033.

[25] Qihang Yao et al. "Multi-class Arrhythmia detection from 12-lead varied-length ECG using Attention-based Time-Incremental Convolutional Neural Network". In: *Inf. Fusion* 53.February 2019 (2020), pp. 174–182. ISSN: 15662535. DOI: 10.1016/j.inffus.2019.06.024.

[26] Özal Yıldırım et al. "Arrhythmia detection using deep convolutional neural network with long duration ECG signals". In: *Computers in biology and medicine* 102 (2018), pp. 411–420.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:54:27 UTC from IEEE Xplore. Restrictions apply.

Computing
https://doi.org/10.1007/s00607-021-00960-8

REGULAR PAPER



MDER: modified degree with exclusion ratio algorithm for influence maximisation in social networks

Sanjay Kumar[1,2] · Dipti Lohia[1] · Darsh Pratap[1] · Ashutosh Krishna[1] · B. S. Panda[2]

Received: 10 May 2020 / Accepted: 13 May 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2021

Abstract

The online social network has become an integral part of our day to life and serves as an excellent platform for sharing ideas, opinions, and products. Influence maximization (IM) is a widely studied topic in the area of social network analysis. The objective of IM is to find influential nodes that can disseminate information to a larger extent in the network. Many local and global centrality measures are proposed to rank the nodes based on their spreading capability with certain limitations. Many proposed algorithms locate the spreaders sharing overlapping regions or are closely placed, which may cause interference in spreading. In this paper, based on the notion of maximum coverage of the information and minimum interference in spreading, we propose a novel semi-local algorithm named as *modified degree centrality with exclusion ratio* to identify influential nodes from diverse locations in the network. We use modified degree centrality by considering neighbours upto 2-hops and introduce the novel idea of exclusion ratio to ensure minimum overlapping between regions influenced by the chosen spreader nodes. Extensive experimental validation using classical informa-

✉ Sanjay Kumar
sanjay.kumar@dtu.ac.in

Dipti Lohia
diptilohia083@gmail.com

Darsh Pratap
darshpratap@gmail.com

Ashutosh Krishna
akrishna286@gmail.com

B. S. Panda
bspanda@maths.iitd.ac.in

1 Department of Computer Science and Engineering, Delhi Technological University, Main Bawana Road, New Delhi 110042, India

2 Computer Science and Application Group, Department of Mathematics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India

Published online: 07 June 2021

tion diffusion model demonstrates that the proposed method delivers better results in comparison to many popular contemporary methods of influence maximization.

Keywords Influence maximization (IM) · Independent cascade (IC) model · Node centrality · Social networks (SNs) · SIR model · Viral marketing

Mathematics Subject Classification 91D30

1 Introduction

In recent years, online social networks (OSNs) have seen tremendous growth in the number of users, which connects these numerous users to perform various online activities. In recent years, OSNs have been the epicenter of numerous research problems like influence maximization, viral marketing, link prediction, community detection, and many more where researchers from different disciplines contribute. These networks provide a great platform to reach a large number of people to promote the idea and products [1,2]. A social network can be interpreted as a graph $G = (V, E)$ where V denote the individuals in this network, while E represents the edge between the users. Here, an edge corresponds to the virtual acquaintance like friendship, follow-followee, co-authorship between the users. The idea of how one individual influences the thoughts of other individuals in the network has gained significant importance. These people having the spreading capability, are often dubbed as influential nodes, and finding such nodes to maximize the information spread in the network is known as influence maximization [3,4]. The problem of influence maximization intends to find k top spreaders from a set of n individuals ($n >> k$) in a system such that information diffusion can be maximized, where k is some constant. The selected influential nodes are often called as seed nodes, which play a significant role in creating an information diffusion cascade to influence many other nodes in the network. Effectively identifying these influential nodes in a network may lead to huge implications in applications such as viral marketing [5]. Viral marketing focuses on targeting these influential nodes for advertisements so as to maximize information dissemination. Many e-commerce companies are utilizing the idea of viral marketing to target many users for successful marketing and promoting their products. The notion behind viral marketing is the word of mouth effect. For instance, people often recommend good restaurants to their friends and family, which increases the probability of them visiting the same restaurant. The complex societal network is highly interdependent, and influence over one individual affects people in their vicinity. Apart from viral marketing, application of influence maximisation extend to rumour control [6], community detection [7], opinion leader detection [8].

The process of influence maximization (IM) usually consists of two phases—identifying influential seed nodes and then the information diffusion phase. Many node centrality based approaches have been introduced to efficiently identify influential nodes like degree centrality [9], closeness centrality [10], eigenvector centrality [11], PageRank [12]. These approaches give importance to each node based on its structural position and its connectivity to other nodes in a network. Most of the centrality based

approaches find influential nodes in a biased manner by just taking into account of the strength of each node in the network. Numerous greedy-based algorithms [13,14] are proposed to solve the influence maximization problem using a greedy strategy and discrete optimization technique. Besides, many community-based influence maximization methods [15,16] are proposed with considerably improved performance. Many of these proposed measures usually come with certain limitations like high time complexity and inappropriate selection of spreader nodes.

In this paper, we propose a novel method to identify influential nodes in a complex network, named as modified degree with exclusion ratio (MDER). The proposed algorithm is a type of semi-local approach, which modifies the degree centrality of each node by considering the neighbors up to 2-hops. We argue that by considering the nodes up to 2-hop may cover the neighborhood of a node adequately for the information spreading process. We introduce the idea of exclusion ratio (ER) to ensure that selected spreaders are from diverse localities. The notion of exclusion ratio prevents the inherent biases for the selection of seed nodes belonging to highly dense regions in a network, usually, which is seen in most of the algorithms. Finally, we calculate the MDER centrality of each node by combining the value of the modified degree and exclusion ratio. The proposed method intends to identify seed nodes from a network such that the regions covered by them have a minimum intersection, i.e., in a particular locality of a network once a higher-ranked node is selected, then the chances of selection of the other higher-ranked node in the nearby locality is decreased. We adapt two popular information diffusion models, namely susceptible-infected-recovered (SIR) model [17] and independent cascade (IC) [18], to measure the spread of the information originating from source nodes. Further, we practice various evaluation criteria to judge the performance of the proposed algorithm, along with other existing methods. The results of the experimental results show that the performance of the proposed algorithm exceeds other existing classical methods. Our major contributions are summarised as follows:

1. We propose a novel approach, modified degree with exclusion ratio (MDER), to identify influential nodes in social networks through the concept of modified degree centrality and mutual exclusion.
2. We evaluate the performance of the proposed approach using ten complex real-life networks of varying applications, size, and complexity.
3. The extensive experimental validation using different performance matrices concludes that the proposed MDER algorithm outperforms many popular methods.

The remaining of the paper is organized as follows: In Sect. 2, we review the literature on various centrality measures, followed by a discussion of diffusion models used for measuring information spread. In Sect. 3, the proposed approach, "modified degree using exclusion ratio," dubbed as MDER, is described in detail with a toy network simulation. In Sect. 4, benchmark datasets and performance metrics used for extensive experimental validation of the proposed approach are discussed. Section 5 summarise the results of the comparison of the proposed approach with existing state-of-the-art techniques. Finally, in Sect. 6, concluding remarks are made.

2 Preliminary

2.1 Influence maximization and viral marketing

Social networks are a type of complex network and are dynamic in nature, where individuals interact with each other and share their interests. This involves a large amount of information flow between the inter-connected entities. A widely studied problem in the field of social network analysis is finding influential nodes in a network to achieve influence maximization. Due to the large and complicated structure of the network, influence maximization is a challenging problem. It is a kind of optimization problem and has mathematically proved that getting an optimal answer is NP-hard using general information spreading models [3]. In social networks, hemophiliac behavior is observed among individuals [19], which means that nodes that are friends or directly connected may exhibit similar behavior and therefore indulge themselves in similar activities and share similar interests. For example, if an individual in a group buys a cellular phone, then the people connected to this individual may be more inclined to buy the same phone as compared to individuals outside this group. This is the central concept of viral marketing: the company targets those individuals who have the maximum reach in the social network and would help them extend their product horizon. Hence, viral marketing thrives on exploiting the power of the influential nodes, which play a significant role in shaping views and opinions of the masses through the word of mouth effect.

In the formulation of influence maximization (IM) algorithm, the main objective is that selected spreaders nodes can propagate the information properly in their neighborhood, leading to a maximum coverage of the information in the network at the end of the diffusion model through cascade triggering. Another crucial factor is that the selection of the spreaders is from diverse regions. Spreaders selected from close to each other may cause interference in the spreading process. Therefore, we aim at maximum coverage and minimum interference or overlapping of spreading caused between selected seed nodes.

2.2 Node centrality

Many node centrality measures have been introduced over the years to rank the nodes in order of their importance in the network. The importance of a node is evaluated based on how a node can spread a piece of information. The centrality measures exploit different topological features of the nodes like the distance between the nodes, shortest path between nodes, the number of connections a node has, or how deeply is a node is lying in the network. The following section describes the various classical centrality measures.

Degree centrality (DC) [9] counts the number of neighbors of a node. There are two types of degree: in-degree and out-degree. The in-degree counts the number of incoming edges while the out-degree counts the number of outgoing edges. The time complexity of computing the following method is $O(m + n)$, where n is the total number of nodes, and m is the number of edges. This is a simple yet effective measure

to find influential nodes. One of the major drawbacks of degree centrality is that it deals with the local structure of the network by considering only the direct edges connected to the node. Betweenness centrality [20] incorporates the global structure of the system. It calculates the number of times a node (u) lies between the shortest path of any two nodes. The idea behind this method is that the flow of information must happen from the node that is part of most of the shortest paths. Since this is a global metric, it has a high time complexity of $O(mn + n^2 logn)$. In Eq. 1, B_u is the betweenness centrality value of each node $u \in V$, $\sigma_{ij}(u)$ is the number of shortest paths that contain node u and σ_{ij} is the total number of paths between each pair of nodes i and j.

$$B_u = \sum_{u!=i!=j \in V} \frac{\sigma_{ij}(u)}{\sigma_{ij}} \tag{1}$$

Closeness centrality (CC) [10] determines how close a node u is to other nodes by calculating the distance between u and other nodes. Like betweenness centrality, this method considers the global structure of the system and therefore takes O(n^3) time in producing results. The Eq. 2 given below calculates the closeness centrality, C_u for each node u where $d(u, i)$ is the distance between nodes u and nodes i.

$$C_u = \frac{1}{\sum_{i=1}^{n} d(u, i)} \tag{2}$$

The basic idea behind Eigenvector centrality (EC) [11] is that a node considered important if it is connected to other important nodes in the system. The time complexity is $O(n^2)$. Google mainly uses PageRank [12] approach for ranking of their web pages. Here, each webpage is assigned a score calculated based on the number of links the web page contains and the weights of the links. The web page is analogous to a node, while the links are the edges. The PageRank approach works best with directed graphs. For instance, Degree centrality takes into account the number of neighbours of each node, thereby neglecting global structure. Although betweenness centrality and closeness centrality are based on global measures and provide satisfactory results, computation complexity in these methods is quite high for large networks. The k-shell centrality [21] approach assigns value to nodes based on how deeply the node is lying in the network. The core nodes or the central nodes get high value as compared to the ones on the peripheral of the network. This is done by pruning the nodes recursively until the graph is empty. The time complexity involved in finding k-shell values for all nodes is $O(m+n)$. However, this approach assigns similar core values to many nodes and does not take into account that some of the nodes having the same core value are more influential. Ma et al. [22] proposed gravity centrality inspired by the classical theory of gravity. They adapt the k-shell value of each node as its mass and the shortest path distance between the source node and all other nodes as their distance. The Eq. 3 calculates the gravity centrality(G_u) for node u. Here, ks_u represents the k-shell value of node u, ks_v represents k-shell value of the node v and d_{uv}^2 is the square of the shortest path distance between node u and node v. To compute the gravity centrality

of a node u, all nodes up to k-hops from u are considered where k denotes the average degree of the nodes.

$$G_u = \sum_{i=1}^{k} \frac{ks_u * ks_v}{d_{uv}^2} \tag{3}$$

The time complexity of gravity centrality is $O((m+n)k^3)$.

Hirsch index or h-index is a popular matrix to assess the impact of a researcher based on the number of citations received. The concept can be extended in node centrality [23]. The h-index centrality of a node is the maximum h-value such that it has at least h neighbors, and each of these neighbors must have a degree greater or equal to the value of h. Sheikhamadi et al. [24] proposed a DegreeDistance approach which improves in two phases, FIDD and SIDD. They recognised the common drawback of many centrality measures that the selected set of seed nodes have a high number of common neighbours and they overcome this by applying the distance measure to select appropriate nodes. Berahmand et al. [25] proposed a local centrality based spreading method named DCL, which ranks the nodes based on its degree, the degree of its neighbors, common connections between the neighbors of the node, and inverse local cluster coefficient. Authors in [26] introduced a local centrality approach based on the notion of edge ratio and neighborhood diversity measure with a focus on identifying global bridge for the maximum spread of the information. They argued that a node with a high edge ratio and a more diverse neighborhood could contribute significantly to information diffusion. Rui et al. [27] suggested a type of local method of IM named Fixed Neighbor Scale (FNS), which considers neighbors of a node up to multiple levels to determine its spreading influence. They determine the spreading capability of a node by adding multi-level neighbors' weights in the form of their distances from the source node.

2.3 Information diffusion

The process of circulation of information to various users originating from some source nodes in a network is called information dissemination or information diffusion. The modeling of information propagation requires a suitable mathematical formulation [28]. Researchers use the various information diffusion models to evaluate how an influence maximization approach is useful in spreading the information in a real-life scenario. One of the popular categories of information diffusion models is based on epidemic spreading [29]. This is based on how a contagious disease spreads over a network. Initially, few people are infected with the disease, and the rest of them are prone to it (i.e., susceptible). The disease can flow from the infected to the susceptible person. In this manner, a piece of information can spread in a network. The susceptible-infected-recovered (SIR) model is one of the widely used information diffusion model [17]. In this model, nodes are classified into three categories: (i) S-susceptible nodes are the ones that are vulnerable to the disease or information. (ii) I-infected nodes are the ones who know the information, and they can spread it to the nodes that are in

its vicinity, and (iii) R-recovered nodes are the ones that were infected and now have been cured. The recovered nodes can not spread the information further.

Other popular categories of information diffusion models are predictive models such as the independent cascade (IC) [18] and the linear threshold (LT) model. In IC model, each inactive node v can be activated by an active adjacent node u with a probability $p(u, v)$. In other words, if at the beginning few of the nodes are active, let's say u is one such node, while the rest are inactive, let's v is one such node, then at time t, u that was activated at time $t-1$ can activate its neighbor v with probability $p(u, v)$. This probability is called the activation ratio. Once the node is activated, it remains activated. LT model [30], on the other hand, is similar to the IC model but is based on the cumulative effect of the active nodes. Each node decides on a threshold value ranging in [0, 1]. At every step, node v check the weighted sum of the active neighbors. If the weighted sum is greater than the threshold, then v gets activated. Once the node is activated, it remains activated. For the evaluation of influence spread, in this manuscript, we use the SIR and IC model.

3 Proposed algorithm: MDER

In this section, we present our proposed algorithm based on the notion of maximum coverage and minimum interference or overlapping of spreading caused between selected seed nodes. Here, overlapping means if a high-rank node (u) can diffuse information to an area independently, then the other influential node (v) in the same locality may cause overlapping in the diffusion process and therefore is kept relatively lower in the rank list. We introduce the idea of exclusion ratio (ER) in our approach to take care of minimum interference in information dissemination. Therefore, our main objective in case of influence maximization is to choose the seed nodes from the different parts or locality of the network rather than choosing multiple high-rank nodes that are in lying in the vicinity of each other. This is especially useful for an e-commerce company to target influential customers of different regions to promote the product through word-of-mouth strategy and to optimize the overall advertisement. The brief overview of the proposed algorithm and its working is depicted in Fig. 1. The proposed algorithm is outlined as follows:

The brief descriptions of Algorithm 1 are: *Step 1* Firstly, we assign a score to each node (u) by slightly modifying degree centrality. For this we consider i-hop neighbours of node u. These neighbours are used to compute modified degree score of u as follows:

$$MD_i(u) = \frac{\text{No. of } i\text{-hop neighbours}}{\text{total no. of nodes in the network}} \tag{4}$$

Next, we rank the nodes in the network based on the descending order modified degree calculated above and store in list l_i.

Step 2 In this step, we try to ensure mutual exclusion. For doing this, we calculate the exclusion ratio (ER) associated with each node (v). We sequentially pick nodes from the list (l_i) generated in step 1. For each node (v) we introduce two parameters,

Algorithm 1 : MDER

Input: $G = (V, E)$, where $n = \|V\|, m = \|E\|$
Output: $\mathcal{R}$
1: **for** $i \in \{1, 2\}$ **do**
2: **for** $v \in \{1, 2, ..., n\}$ **do**
3: $k =$ number of i-hop neighbours of node u
4: $MD_i(v) = \frac{k}{n}$
5: **end for**
6: $l_i \leftarrow$ sorted list of V in descending order of $MD_i(u)$
7: $MDER_i \leftarrow \{\}$
8: **for** each vertex $v \in l$ **do**
9: $old_v =$ number of i-hop neighbours of v that are also i-hop neighbours of the nodes higher than v in the list l_i
10: $new_v =$ number of i-hop neighbours of v that are not included in old_v
11: Calculate the value of $ER(v)$ for each node v using Eq. No. 5.
12: $MDER_i(v) = MD_i(v) \times ER(v)$
13: **end for**
14: **end for**
15: **for** $v \in \{1, 2, ..., n\}$ **do**
16: $MDER(v) = MDER_1(v) \times MDER_2(v)$
17: **end for**
18: $\mathcal{R}$ = top c nodes based on the final $MDER$ score.
19: **return** $\mathcal{R}$

namely old and new, where, v_{new} is i-hop neighbors of v that have not been previously covered by any higher-ranked node other than v, while v_{old} counts i-hop neighbors of v that have already been covered by a higher-ranked node present in list l_i. Now, we compute the exclusion ratio (ER) as follows:

The exclusion ratio ensures a minimum overlapping between regions influenced by the nodes. Hence, the notion of exclusion ratio ensures that selected influential nodes lead to maximize information dissemination and prevents the inherent biases for nodes belonging to highly dense regions in a network, which tend to have a higher probability of selection.

Step 3 Now, we compute $MDER_i$ score for each node v by considering its i-hop neighbors. The initially computed centrality value MD in step 1, is multiplied by the exclusion ratio ER for that node, computed in step 2. Hence, the final centrality value of each node v is defined as:

$$MDER_i(v) = MD_i(v) \times ER_i(v) \tag{5}$$

Step number 1, 2 and 3 are going to execute two times for computing MDER score of each node (v), as mentioned in the line no. 1 of the proposed Algorithm 1. First time for $i = 1$ that considers 1-hop neighbors and second time for $i = 2$ that considers 2-hop neighbors.

Step 4 Finally, we combine MDER results computed for 1-hop neighbours as well as 2-hop neighbours i.e $i = 1$ and $i = 2$ to get the final MDER value for each node v by using the mathematical relation:

$$MDER(v) = MDER_{i=1}(v) \times MDER_{i=2}(v) \tag{6}$$

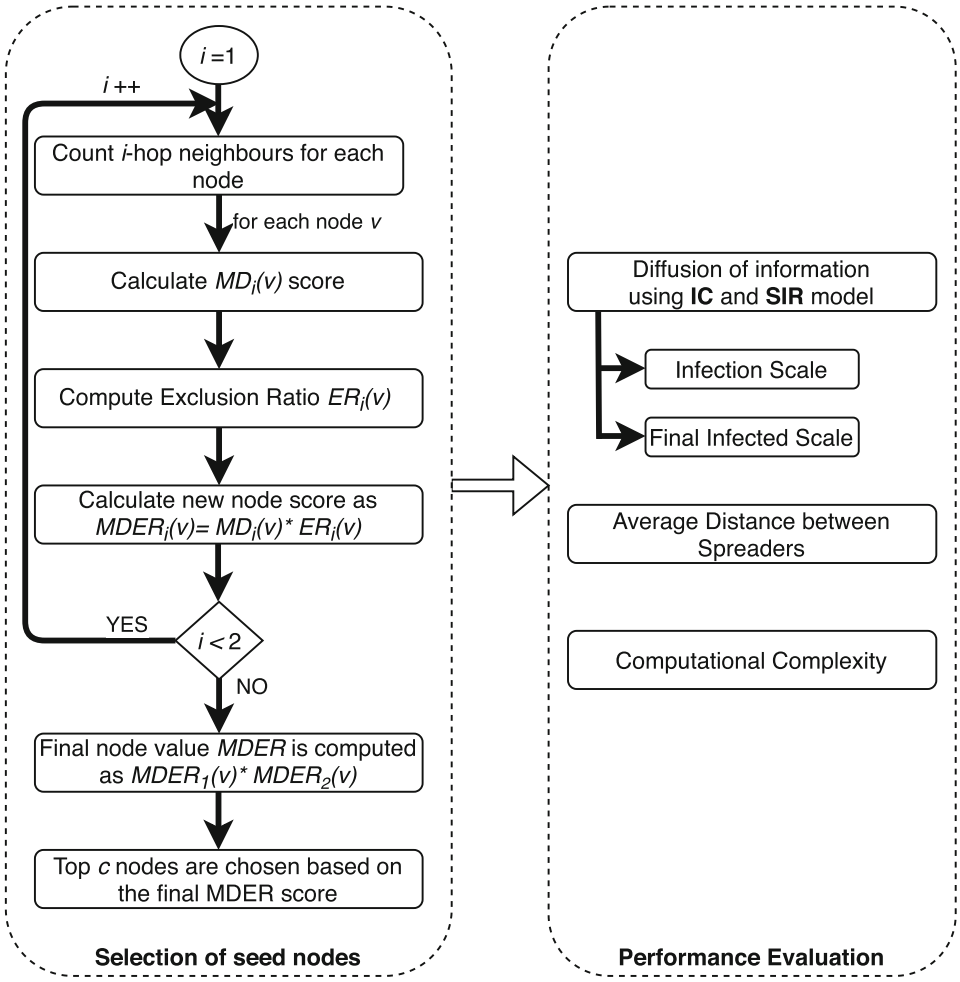


Fig. 1 Brief overview and working of the proposed algorithm

The values in the $MDER_i$ list are normalized before multiplying so that one of the factors does not overpower the other and does not reduce the effect of each other completely.

We argue that in case of influence maximization, mostly a node can influence its neighbors up to 2-hops. This implies that by considering all the nodes up to 2-hops from a source node u, its neighborhood can be adequately covered in terms of spreading influence. For this, we consider the various combination of $i = 1, 2, 3, 4$-hop neighbors, and practically we achieved the best performance in the influence spread by combining 1-hop and 2-hop neighbors. The comparison of the MDER value for different values of i ($i \in 1, 2, 3, 4$) has been made as shown in the Fig. 2. The results in the form of the final infected scale versus spreaders fractions using the SIR information diffusion model on gplus, facebook, PGP and ca-Hepth. Here, Ca-hepth is a collaboration network dataset, gplus, and facebook belong to the social network dataset, and PGP is a communication network dataset. The obtained results show that the proposed MDER (in blue color), which is the hybrid of the values of $MDER_{i=1}$ and $MDER_{i=2}$ outperforms the others where $MDER-1$, $MDER-2$, $MDER-3$ and $MDER-4$ represents MDER score of each node by considering only its 1-hop, 2-hop, 3-hop, and 4-hop neighbors respectively.

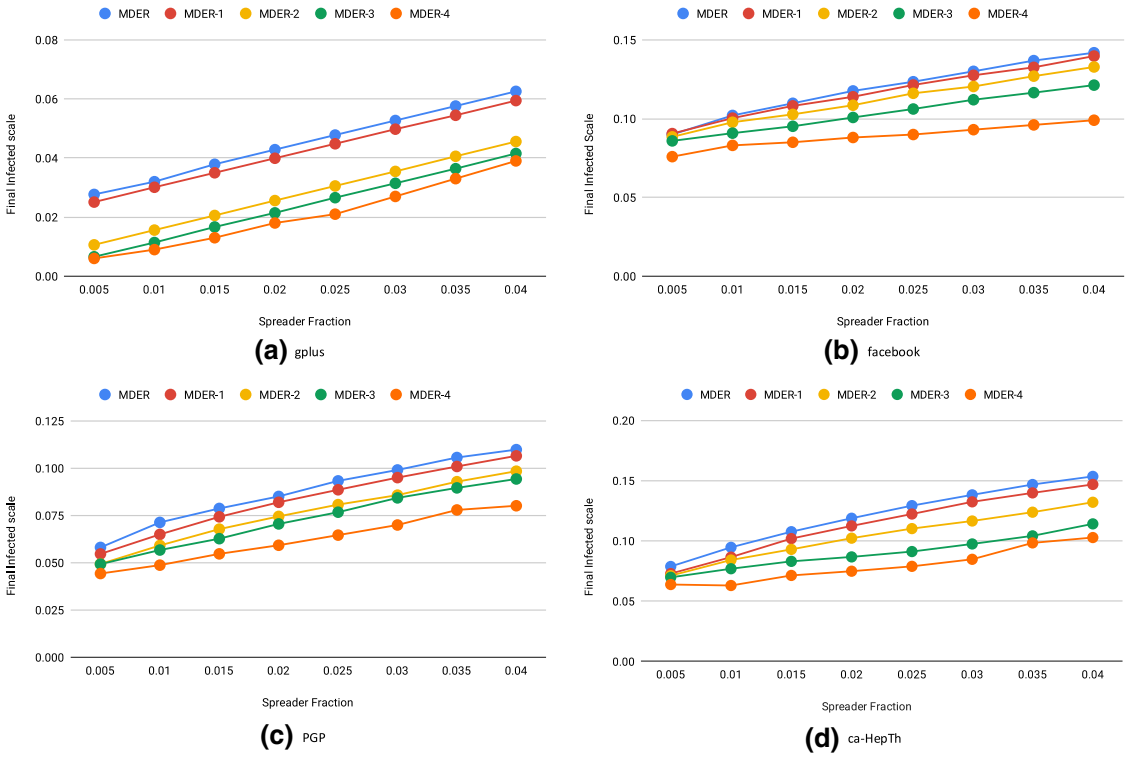


Fig. 2 Final infected scale versus spraeder fractions using SIR model for **a** gplus, **b** facebook, **c** PGP, and **d** ca-Hepth dataset. The result shows the comparision of results for different values of $MDER_i$ (color figure online)

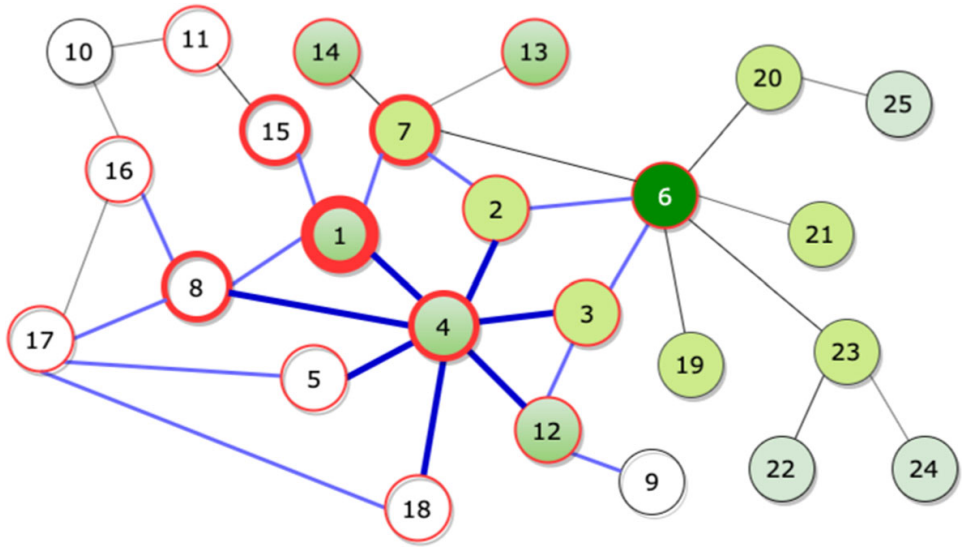


Fig. 3 Toy Network having 25 nodes. The neighbours of node number 6 upto 2nd level is represented in green color. The neighbours of node number 1 upto 2^{nd} level is outlined with red color and the neighbours of node number 4 upto 2nd level is shown having edges of the color blue (color figure online)

3.1 Explanation on how our algorithm achieves mutual exclusion through a toy network

The proposed algorithm is explained with the help of a toy network consisting of 25 nodes in Fig. 3. The node numbered 6 (marked in green color) has got the highest

MDER score, followed by node numbered 1 (marked in red color). The neighbors of these nodes are marked with a lighter version of their respective colors.

The idea of exclusion ratio is illustrated in the toy network is given as follows: Node number 6 ranks the highest in list, this means that the value of parameter old for node 6 is zero, and the value of parameter new for node 6 is 15, which is total number of neighbours up to 2-hops. Hence, the value of $E(6) = 15$. Second in the rank list is node numbered 1. For node 1, the value of parameter old is the number of common 2-hop neighbours of nodes 6 and 1, since node 6 is higher in the list. The list of common neighbours up to 2-hop of both nodes includes node number 2, 3, 4, 6, 7, 12, 13, and 14. Therefore, the value of parameter old for node 1 is 8 and the value of parameter new for node 1 is the number of its remaining neighbours up to 2-hops, which is equal to 7 and includes node numbered 5, 8, 11, 15, 16, 17, and 18. Hence the value of exclusion ratio for node 1, i.e. $E(1)= \frac{7}{8}$.

According to degree centrality, node 4 and node 6 should be in the same position in the rank list since the number of neighbors for them is the same, i.e., 7. But, the MDER method ranks node number 4 lower in the rank list. According to our method, nodes in the neighborhood of nodes 6 and 1 completely cover the nodes covered by 4. This means that the exclusion ratio, for node 4, is very low since it fails to cover multiple new nodes. Therefore, node 4 is not given a higher position in the rank list. This ensures that there is minimal overlap between the regions covered by higher-ranked nodes hence maximizing the diffusion of information.

3.2 Computational complexity

The time complexity of the proposed approach is determined from the pseudo-code presented in Algorithm 1. The outer for loop in line number 1 runs only two times, as we consider neighbors up to 2-hops. In lines from 2 to 5, where the list $MD(v)$ is calculated, the time taken is $O(n + m)$. In line number 6, to rearrange the list in descending order of values can be obtained using a general sorting algorithm like heap sort or quick sort in time $O(nlog(n))$. To calculate the exclusion ratio (ER) and MDER score for each node (v) in lines 8 to 13, requires $O(m + n)$ time, as for each node, we have to look up its neighbors up to 2-hops. In line number 18, we need to find top c nodes, as in influence maximization (IM) objective is to find constant seed nodes, where $c << n$. In the worst case, it can be done in $O(c * n)$ time. Therefore, the final time complexity of the proposed algorithm is $O(n + m) + O(nlog(n)) + O(n * c) = O(nlog(n))$ in the worst case. Here, we considered the time involved in sorting activity in line number 6 as $O(nlog(n))$. But in the actual scenario, we can ignore those nodes in the sorting activity whose number of neighbors up to 2-hops are very less, as in case of the real-life networks preferential attachment and power-law is observed in the degree distribution, which implies many nodes in the network has very few degrees and very few nodes have relatively high degree [31]. By doing some prepossessing on the dataset, we can discard those nodes in the ranking whose degree is relatively very less. Thus the actual running time of MDER is going to be far less than $O(nlogn)$, which is further evident from the results obtained in Sect. 5.4.

4 Dataset and performance metrics

In order to perform the detailed performance analysis of the proposed algorithm, along with popular existing algorithms of IM, we use ten real-life datasets of various sizes and nature and adopt four different evaluation criteria. The datasets and evaluation metrics used are presented below. Most of the practiced real-life datasets are referred from SNAP Stanford [32], which is a publicly available repository maintained by the researchers of Stanford.

4.1 Dataset

In this paper, we utilize a total of ten datasets of varying application, size, nature. Out of that, there are six social network datasets, namely twitter, gplus, twitter-ego, Brightkite, musae-facebook, and musae-twitch. There are three collaboration networks: Enron-email, ca-Condmat, and ca-Hepth and one communication network, Pretty Good Privacy (PGP). Table 1 lists the statistical properties of each dataset where d_{avg} and d_{max} represent the average degree and maximum degree of nodes, respectively. The β_{th} denotes the rate of infection for the information spread. We need to suitably choose the rate of infection (β_{th}) in SIR and activation probability (p_{uv}) in the IC model, which decides the rate at which an infected or active node infects or influences its neighbors. In the literature, most of the IM algorithms have considered the rate of infection as an epidemic threshold which is calculated as

$$\beta_{th} = \frac{\sum_{i=1}^{n} d_i}{\sum_{i=1}^{n} (d_i)^2}, \tag{7}$$

where d_i denote the degree of node v_i. Hence, the value β_{th} for a dataset is equal to the average degree of the nodes divided by the square of the average degree of the nodes.

Enron-email consists of communication of half a million emails among the employees of the Enron corporation [33]. If an email is sent from an employee to another, an undirected link is made between the two. The CA-Condmat dataset consists of nodes as authors, and an edge between the nodes exists if a physics author has co-authored with another author on a paper [34]. Musae-facebook network contains page-page networks of verified Facebook sites. Nodes represent official Facebook pages, and the links are the mutual likes between them [35]. PGP, i.e., Pretty Good Privacy, is an encrypted communication network [36]. Ca-Hepth is a collaboration network of authors who collaborated on papers that were submitted to the High Energy Physics—Theory category [34]. Musae-twitch is a social network dataset, where the nodes are the twitch users, and the link is between them if they are friends online [35]. Brightkite is a location based social network that logs your location and shares it with your connections. This dataset is obtained from Brightkite social network containing friendship networks [37]. Gplus is a social network dataset consists of 'circles' from Google+. The data was obtained from users who had manually shared their circles using the 'share circle' feature in Google+ [38]. Twitter is a widely used social networking ser-

Table 1 Brief statistical descriptions of the different datasets

S.no.	Dataset	No. of nodes (n)	No. of edges (m)	d_{max}	d_{avg}	β_{th}
1	Twitter	81,306	1,342,296	3383	33.018	0.00579
2	Brightkite	58,228	214,078	1134	7.353	0.01569
3	Enron-email	36,697	183,835	1383	10.020	0.00713
4	Gplus	23,628	39,194	2761	3.317	0.00265
5	Twitter-ego	23,370	32,831	238	2.809	0.02597
6	Ca-Condmat	23,133	93,439	279	8.078	0.04533
7	Musae-facebook	22,470	171,002	709	15.204	0.0163
8	PGP	10,638	24,301	205	4.568	0.05299
9	Ca-hepth	9882	25,977	65	5.257	0.07984
10	Musae-twitch	7126	35,324	720	9.914	0.01678

vice where individuals interact by sharing short messages or "tweets". A link is formed if they reply, retweet or follow another individual. The Twitter dataset consists of user lists crawled from Twitter. The dataset includes node features, circles, and ego networks [38]. Twitter-ego is a directed networks containing Twitter user-user following information. It is an ego network that has a focal point called an ego and each ego is connected to its social links called as alters. This data is publicly available at http://konect.cc/networks/ego-twitter/.

4.2 Performance matrices

(i) *Infection scale* The infection scale is the measure that indicates the scale at which nodes get infected or become active and then recovered or inactive due to infection originated from selected seed nodes with respect to time. The calculation of the number of nodes that are affected by the information in the spreading process is determined using an information diffusion model. In the SIR model, we select initial seed nodes that start the diffusion of information in the network. These seed nodes pass the information to their adjacent nodes with an infection probability of β and then recover with a probability of δ. The recovered nodes do not further act as seed nodes and therefore stop infecting their neighbors. The Infection scale is the summation of infected nodes and recovered nodes at a particular timestamp. The infection scale initially rises significantly since the number of recovered nodes is very less, and later the process slows down due to the high number of recovered nodes.

(ii) *Influence spread of total number of infected nodes* We calculate the final infection scale, which counts the total number of infected nodes after the end of the information dissemination process originating from selected influential nodes. As discussed SIR model passes information with the rate β and recovers with a rate of δ, whereas in the case of the IC model, the nodes that seed nodes are initially activated with information disperse information to its neighbors and do not recover and become inactive, instead they continue to diffuse information. The rate at which the node is activated is determined using activation probability p_{uv}. Active nodes or the final infected scale includes the total number of nodes that received information at the end of the spreading process.

(iii) *Average distance between spreaders* This metric calculates the average shortest path distance between the selected seed nodes. A large value of the average shortest distance between spreaders implies that the seed nodes are located diversely and not in the vicinity of each other. The selection from diverse regions means that information dissemination is maximized. Assume shortest path distance between two spreader nodes v_i and v_j is d_{ij} and number of chosen spreaders are c. Then the following equation provides the average distance between spreaders (Ls):

$$Ls = \frac{\sum d_{ij}}{\frac{c(c-1)}{2}} \tag{8}$$

where $\sum d_{ij}$ denotes the sum of shortest path between each pair of seed nodes.

(iv) *Time* We also assess the absolute time required to produce the ranking of nodes in terms of their spreading capability by the proposed model and other methods in comparison. The time complexity to produce a ranking list of nodes in the case of degree centrality, k-shell, and h-index is $O(m+n)$. In the case of PageRank and eigenvector centrality, it is equal to $O(n^2)$, and for gravity centrality time complexity is $O((m+n)k^3)$, where m, n, and k represent the number of edges, number of nodes, and the average degree of the nodes in the network respectively. However, in the worst case and without any data prepossessing, the time complexity of the proposed algorithm is $O(nlogn)$. The actual running time on real-life networks can give a better perspective on the effectiveness of the algorithm.

5 Experimental results and analysis

In this section, we present the exclusive experimental results and analysis to judge the performance of the proposed approach. The performance matrices described in Sect. 4.2 are used to obtain comparisons between the proposed algorithm and several baseline measures for influence maximization. These measures include degree centrality (DC) [9], h-index (HI) [23], Eigenvector (EVC) [11], Pagerank (PR) [12], k-shell (KC) [21], and gravity centrality (GC) [22]. To capture the utility of the introduced algorithm, MDER, simulations are performed on ten diverse real-life networks, as depicted in Table 1. To spread the information originating from seed nodes, the value of infection rate or activation probability for each dataset is taken as β_{th}, which is mentioned in Table 1.

The selection of the node to be infected is chosen randomly from the set of neighbors of the seed node. For example, in the case of the PGP dataset, $\beta_{th} = 0.05299$, which implies an active or infected node at time t can randomly infect 5.29% of its neighbors at time $t+1$. We performed the execution to obtain the results on a personal computer with a configuration of 1.8 GHz Intel Core i5 processor and 8 GB of primary memory.

5.1 Infection scale

Figure 4 depicts the result of the infection scale at increasing timestamps of the proposed algorithm, along with other existing methods. Infection scale at any time t is the sum of infected and recovered nodes. For the experiment, we initially activate 100 top spreaders as selected by each algorithm to commence the spreading process. We utilize the SIR model for the information spreading, and the results are averaged over 100 independent simulations. For each dataset, the value of the infection rate is taken as β_{th}. The number of infected nodes is initially increasing, and after a while, when no further infected nodes are present, then the infected scale becomes constant. In the case of ca-CondMat network, we observe that till timestamp 22, the infection is increasing, and after that, the number of infected nodes became constant with a value of 1540. In contrast, other approaches, not only reach recovery state early but also infect less number of nodes. For example, degree centrality infects a total of 1413 nodes and reaches recovery state at timestamp 18. From the results, it is clear that

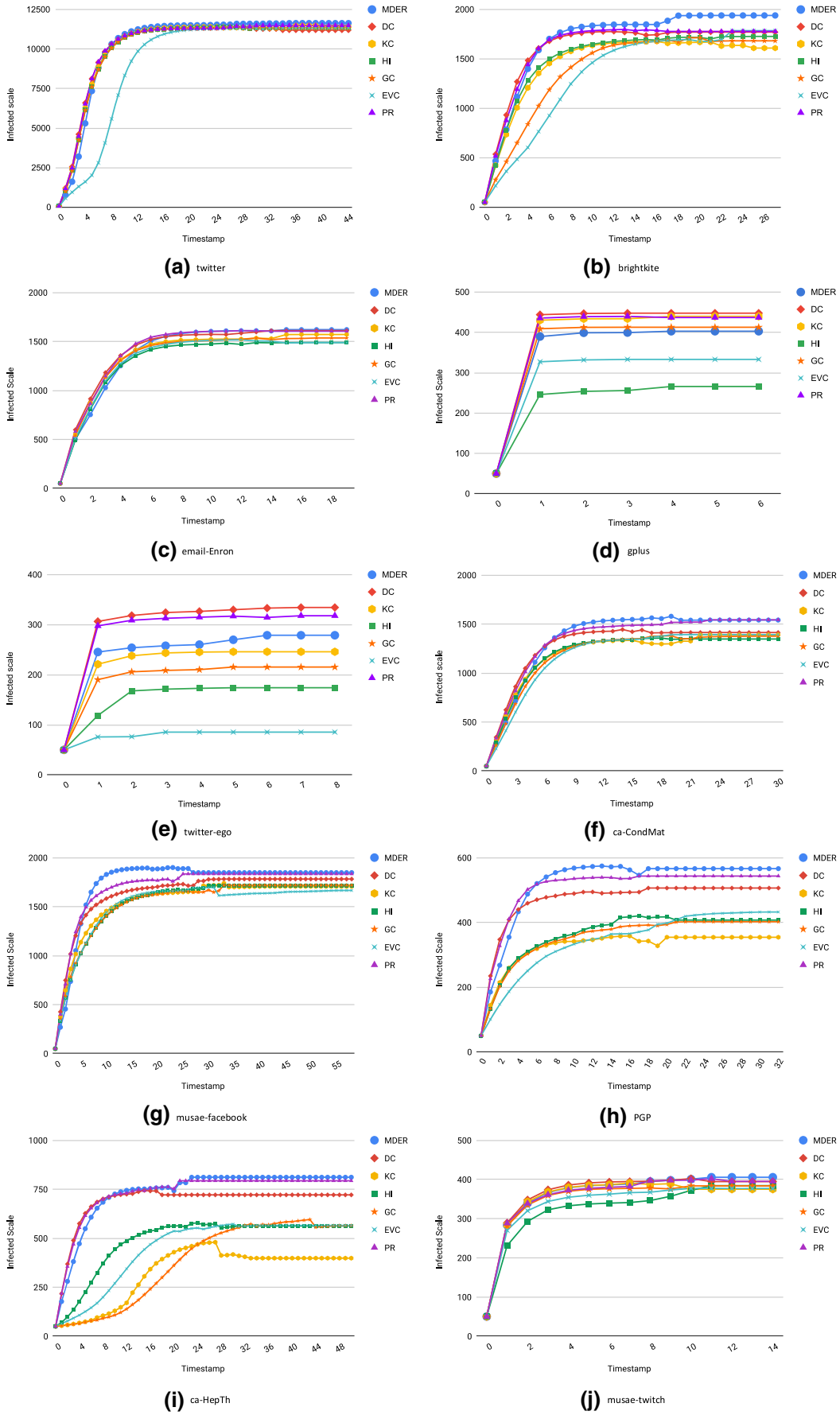


Fig. 4 Infection scale at different timestamps for each dataset by taking top-100 influential nodes from each method. The results are obtained over 100 independent SIR simulations with rate of infection as β_{th}

the proposed MDER algorithm performs better than most of the other approaches. However, in the case of email-Enron and musae-twitch networks, the performance of MDER and Pagerank are very close to each other.

5.2 Final infected scale or active nodes

The final infected scale or active nodes is one of the primary criteria to asses the influence spread produced by triggering from chosen influential nodes. It counts the total number of nodes who got infected or influenced after the end of the spreading process. We utilize the SIR and IC model as the information diffusion model to find the final infected scale with respect to spreader fractions, where spreaders are chosen by MDER and other methods in comparison. To obtain the results for each dataset considered, we keep the spreader fractions in the range of 0.005, 0.01, 0.015, 0.020, 0.025, 0.03, 0.035, and 0.040, and the value of the final infected scale is computed against each fraction of spreader. For example, in the case of the PGP dataset containing 10, 638 nodes, an absolute number of influential spreaders are taken in the range of 53, 106, 159, 212, 265, 319, 372, and 425 to plot the final infected scale.

Figure 5 shows the results of the final infected scale or active nodes versus fraction of influential nodes chosen by MDER and other methods in comparison using the SIR information diffusion model. The obtained results are averaged over 100 simulations of the SIR model with infection rate as β_{th}, which is mentioned in Table 1. In Fig. 5f, for the ca-Condmat network, there are a total of 23, 133 nodes out of which we select 0.005 percent of nodes (which is approximately 115 nodes). We observe that the final infection scale is 1781 nodes. From the results in Fig. 5, it is clear that MDER produces consistent and best results as compared to all other methods in comparison for each dataset. Similarly, Fig. 6 shows the results of total active nodes produced at the end of the spreading process due to influential seed nodes under the IC model. The results are sampled over 100 independent simulations of the IC model. In Fig. 6f for the ca-Condmat dataset, we note that when a fraction of 0.005 nodes are activated initially, then there are a total of 8602 active nodes at the end of the process. From the results in Fig. 6, it is evident that the proposed algorithm beats all other methods comfortably in each dataset.

The performance obtained by MDER using both diffusion models, namely SIR and IC, exceed the performance of other popular measures, including k-shell, PageRank, and gravity centrality. Hence, the experimental results achieved in this matrix echoes the utility and superiority of the quality of seed selection by the proposed algorithm, MDER.

5.3 Average distance between spreaders

To achieve maximum coverage of information with minimum interference cause in spreading, the average shortest distance between the spreaders (L_s) is an important matrix. The value of L_s is computed using Eq. 8 to demonstrate that the selection of nodes is made from diverse regions. The results of this experiment are shown in Fig. 7.

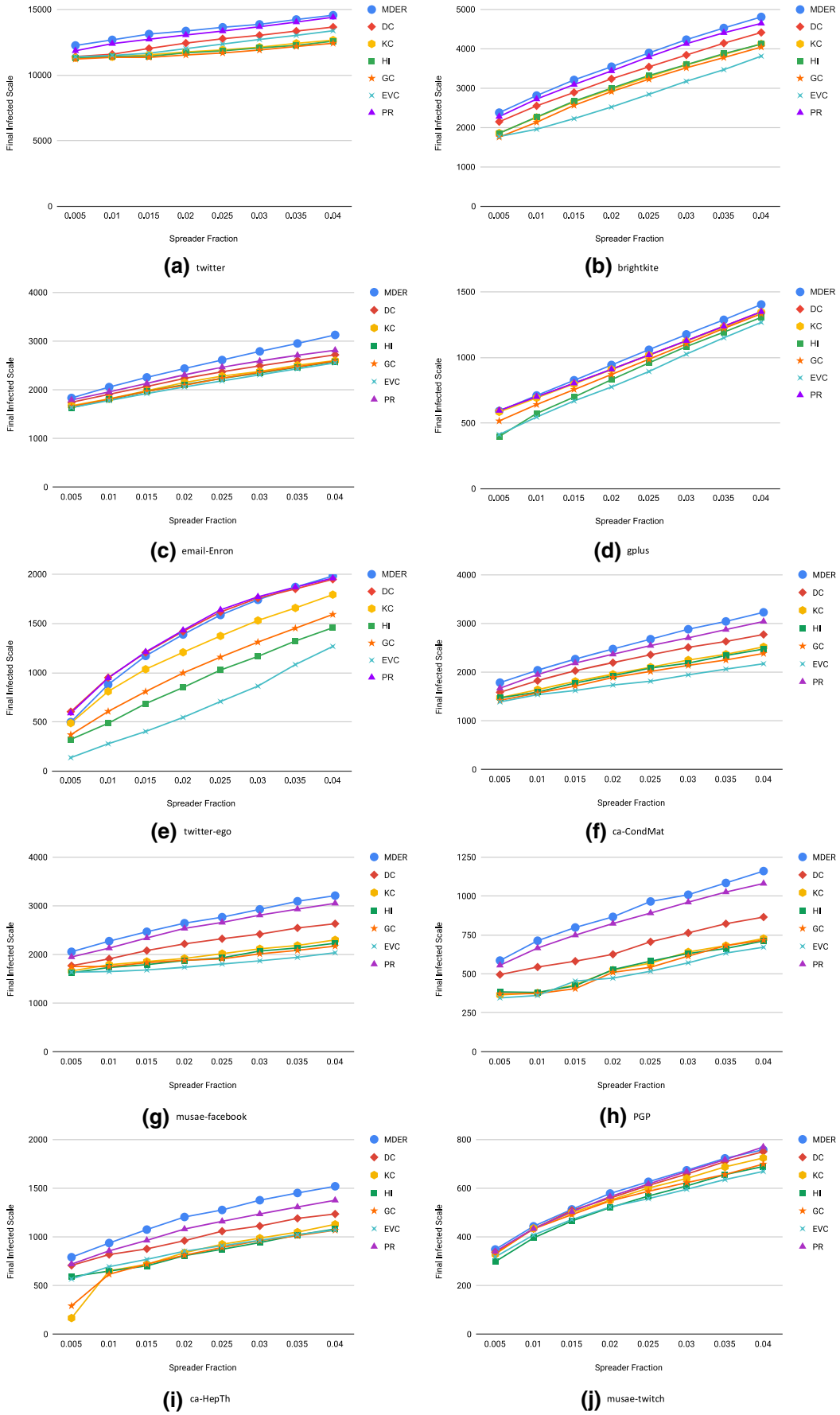


Fig. 5 Final infected scale with respect to spreader fractions using SIR model on ten different real-life networks: **a** twitter, **b** brightkite, **c** email-Enron, **d** gplus, **e** twitter-ego, **f** ca-CondMat, **g** musae-facebook, **h** PGP, **i** ca-HepTh, and **j** musae-twitch. The results are obtained using 100 independent simulations

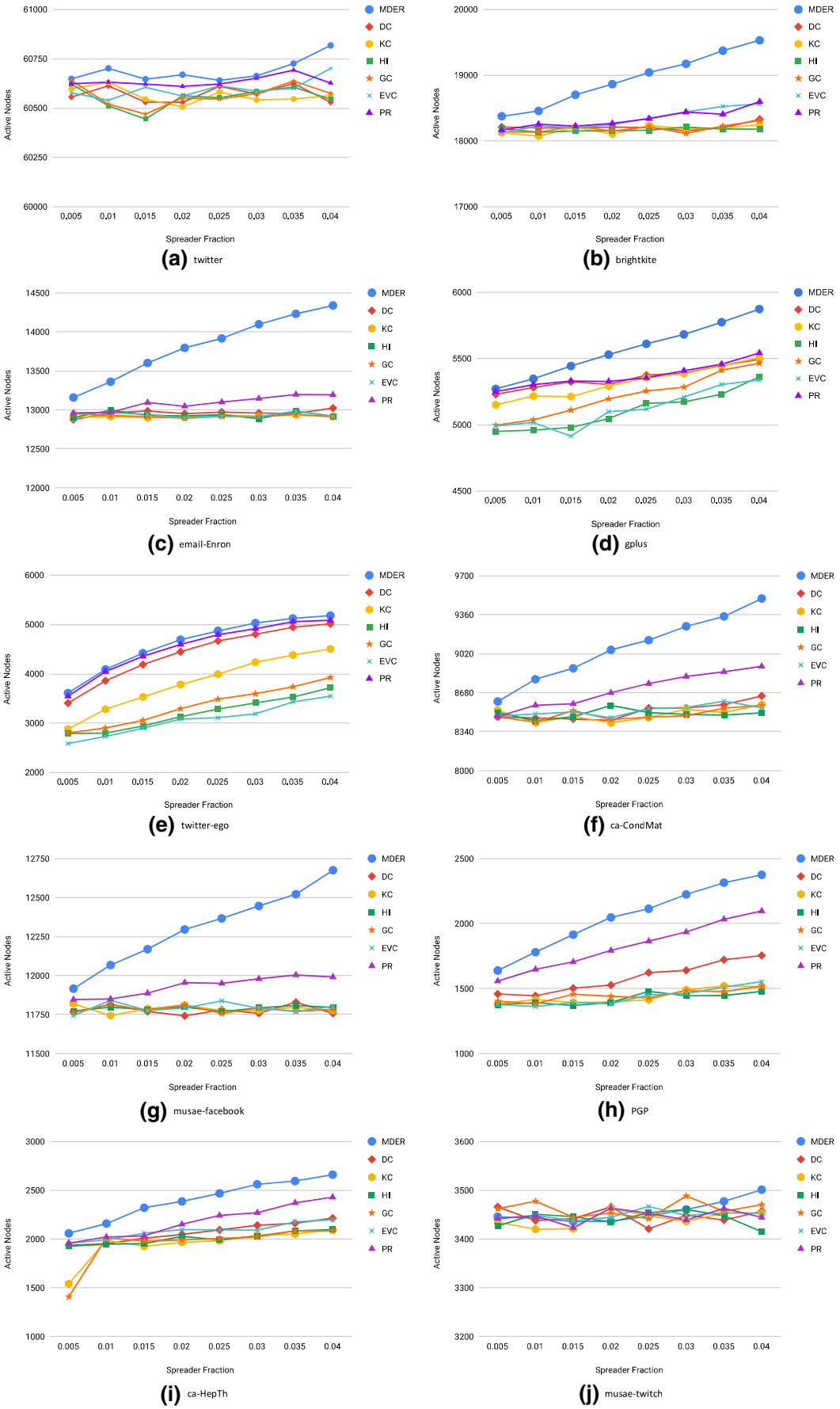


Fig. 6 Number of active nodes with respect to spreader fraction using IC model on six different real networks—**a** twitter, **b** brightkite, **c** email-Enron, **d** gplus, **e** twitter-ego, **f** ca-CondMat, **g** musae-facebook, **h** PGP, **i** ca-HepTh, and **j** musae-twitch. The results are sampled over 100 independent simulations

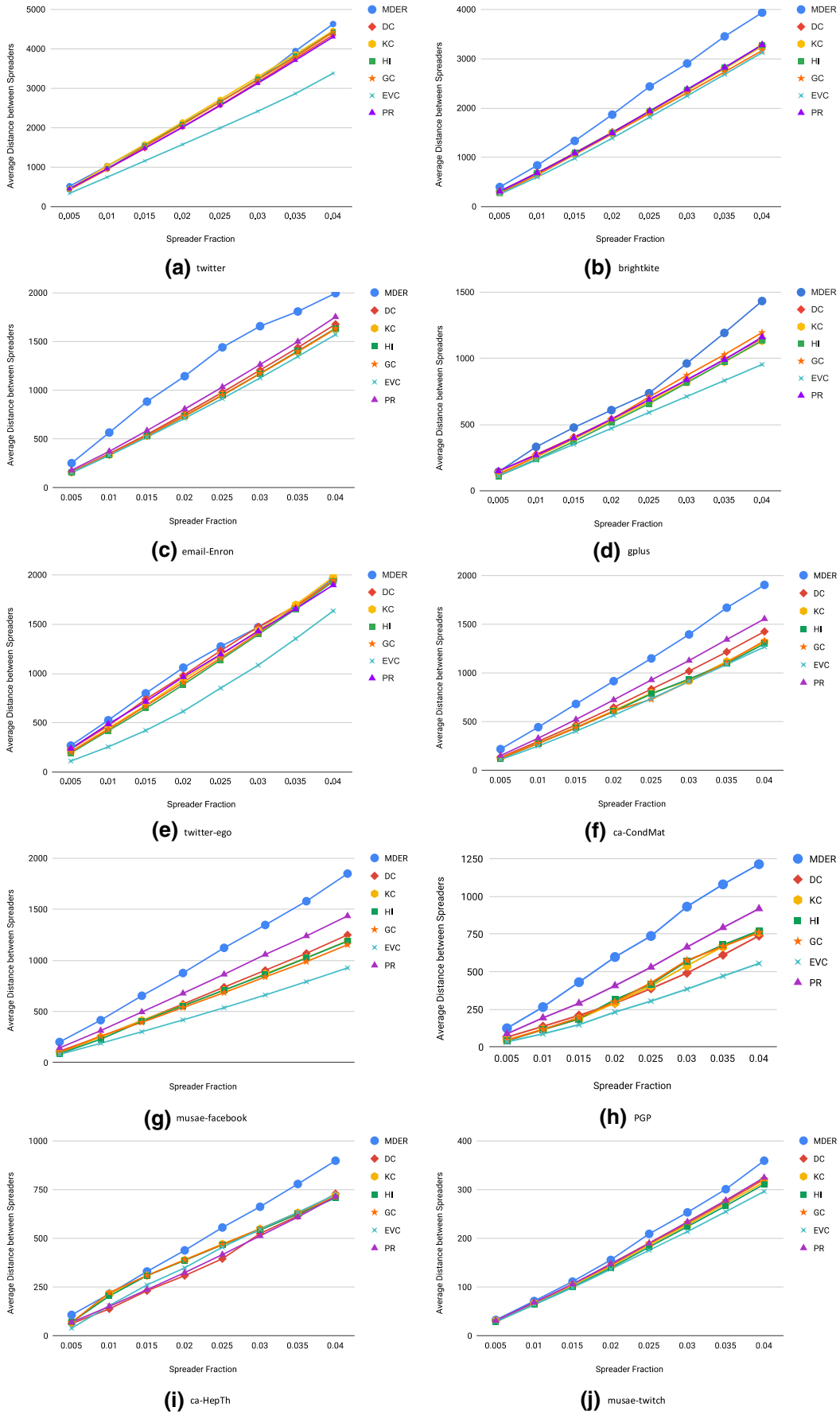


Fig. 7 **a–j** The average shortest path distance between the selected spreaders (L_s) versus spreader fraction for each dataset

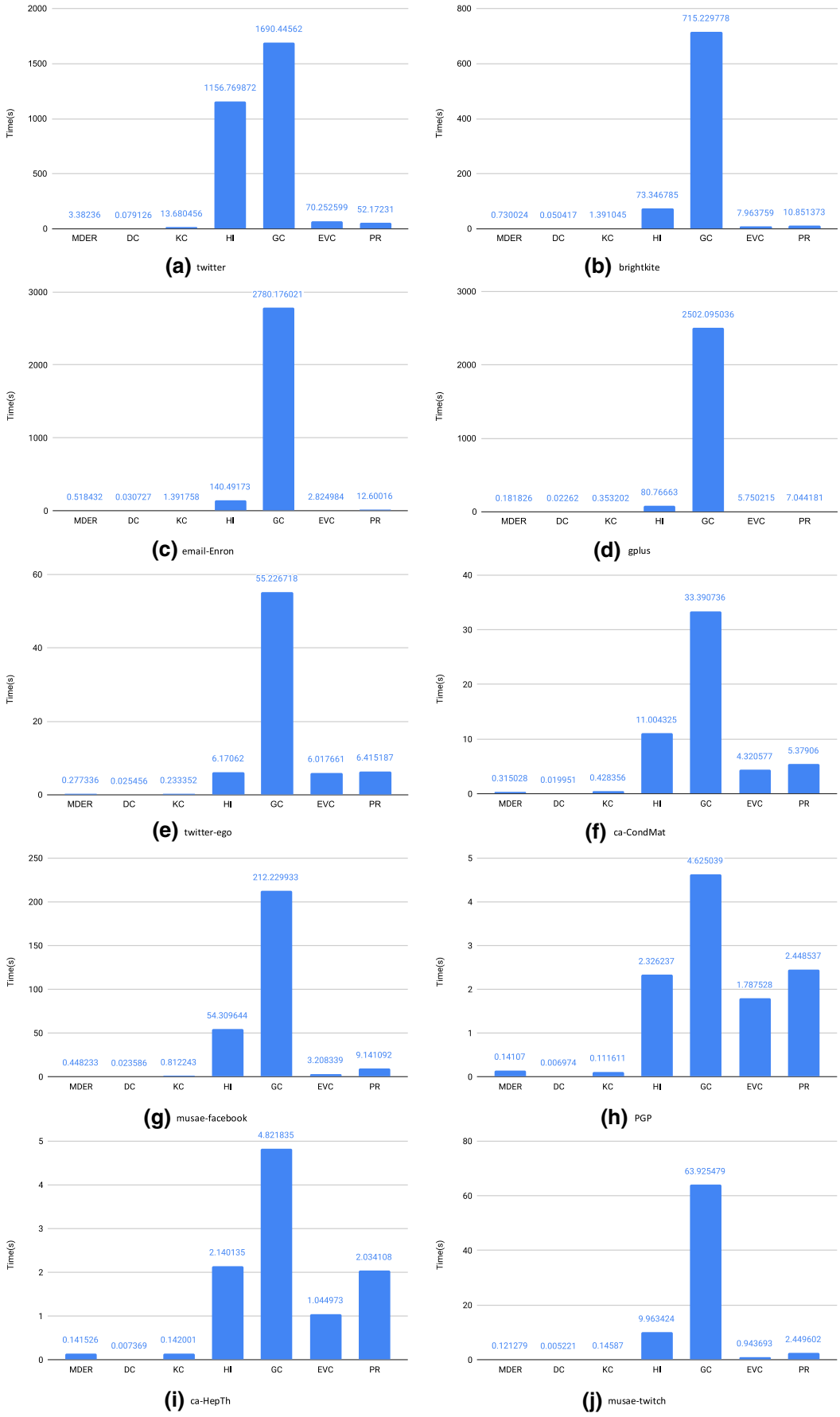


Fig. 8 **a**–**j** The actual running time to generate the ranking list of nodes under different methods for **a** twitter, **b** brightkite, **c** email-Enron, **d** gplus, **e** twitter-ego, **f** ca-CondMat, **g** musae-facebook, **h** PGP, **i** ca-HepTh, and **j** musae-twitch datasets

The graphs show that the distance between the spreaders chosen through the MDER approach is consistently higher than the spreaders elected by other approaches in every dataset. For instance, In the email-Enron network, when 0.04 percent spreaders are selected, then the value of L_s for MDER is 1993, whereas others lie in the range of 1568 and 1749. From the obtained results, it is evident that the proposed algorithm outperforms all other methods.

5.4 Time

In Fig. 8, we compare the actual running time required to rank the nodes based on the spreading capability using the proposed algorithm, MDER, and all other algorithms in the comparison. We performed this examination on ten data sets, namely twitter, brightkite, email-Enron, gplus, twitter-ego, ca-Condmat, musae-facebook, PGP, and ca-HepTh and musae-twitch dataset. We observed that the time taken by MDER for the generation of the ranked list is close to that of degree and k-shell centrality. To process 23, 133 nodes in ca-CondMat, MDER only takes only 0.31 s, which is significantly faster than h-index, gravity centrality, Eigenvector centrality, and PageRank.

6 Conclusion

Finding influential speeders is of utmost importance to achieve optimal advertisement of the products and ideas in a complex network like social networks. In this paper, we introduced an influence maximization algorithm named Modified Degree with Exclusion Ratio (MDER). The method aims to select nodes that have the highest connections based on modified degree centrality up to 2-hops, and the regions covered by the chosen nodes do not overlap with each other. The introduction of mutual exclusion in the proposed approach prevents the inherent biases for the selection of many nodes belonging to highly dense regions in a network. It ensures that influential nodes are selected from diverse regions, which eventually leads to a maximum coverage of the network. The simulation of the proposed algorithm using classical information diffusion models on ten datasets of various sizes and applications echoes the efficiency and utility of the algorithm in real-life scenarios. The method outperforms the widely used centrality measures, including k-shell, PageRank, and gravity centrality.

References

1. Heidemann J, Klier M, Probst F (2012) Online social networks: a survey of a global phenomenon. Comput Netw 56(18):3866–3878
2. Krasnova H, Spiekermann S, Koroleva K, Hildebrand T (2010) Online social networks: why we disclose. J Inf Technol 25(2):109–125
3. Kempe D, Kleinberg J, Tardos É (2003) Maximizing the spread of influence through a social network. In: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, pp 137–146
4. Li Y, Fan J, Wang Y, Tan KL (2018) Influence maximization on social graphs: a survey. IEEE Trans Knowl Data Eng 30(10):1852–72

5. Domingos P, Richardson M (2001) Mining the network value of customers. In: Proceedings of the seventh ACM SIGKDD international conference on knowledge discovery and data mining. ACM, pp 57–66
6. Vega-Oliveros DA, da Fontoura CL, Rodrigues FA (2020) Influence maximization by rumor spreading on correlated networks through community identification. Commun Nonlinear Sci Numer Simul 83:105094
7. Kumar S, Panda BS, Aggarwal D (2020) Community detection in complex networks using network embedding and gravitational search algorithm. J Intell Inf Syst. https://doi.org/10.1007/s10844-020-00625-6
8. Oueslati W, Arrami S, Dhouioui Z, Massaabi M (2021) Opinion leaders' detection in dynamic social networks. Concurr Comput Pract Exp 33(1):e5692. https://doi.org/10.1002/cpe.5692
9. Freeman LC (1978) Centrality in social networks conceptual clarification. Soc Netw 1(3):215–39
10. Okamoto K, Chen W, Li XY (2008) Ranking of closeness centrality for large-scale social networks. In: International workshop on frontiers in algorithmics. Springer, Berlin, Heidelberg, pp 186–195
11. Bonacich P (2007) Some unique properties of eigenvector centrality. Soc Netw 29(4):555–64
12. Brin S, Page L (2012) Reprint of: The anatomy of a large-scale hypertextual web search engine. Comput Netw 56(18):3825–33
13. Cheng S, Shen H, Huang J, Zhang G, Cheng X (2013) Staticgreedy: solving the scalability–accuracy dilemma in influence maximization. In: Proceedings of the 22nd ACM international conference on information & knowledge management, pp 509–518
14. Goyal A, Lu W, Lakshmanan LV (2011) Celf++ optimizing the greedy algorithm for influence maximization in social networks. In: Proceedings of the 20th international conference companion on World wide web, pp 47–48
15. Huang H, Shen H, Meng Z (2020) Community-based influence maximization in attributed networks. Appl Intell 50(2):354–64
16. Kumar S, Singhla L, Jindal K, Grover K, Panda BS (2021) IM-ELPR: Influence maximization in social networks using label propagation based community structure. Appl Intell. https://doi.org/10.1007/s10489-021-02266-w
17. Satsuma J, Willox R, Ramani A, Grammaticos B, Carstea AS (2004) Extending the SIR epidemic model. Phys A Stat Mech Appl 336(3–4):369–75
18. Goldenberg J, Libai B, Muller E (2001) Using complex systems analysis to advance marketing theory development: modeling heterogeneity effects on new product growth through stochastic cellular automata. Acad Mark Sci Rev 9(3):1–8
19. Murase Y, Jo HH, Török J, Kertész J, Kaski K (2019) Structural transition in social networks: the role of homophily. Sci Rep 9(1):1–8
20. Freeman LC (1977) A set of measures of centrality based on betweenness. Sociometry 35–41. https://doi.org/10.2307/3033543
21. Kitsak M, Gallos LK, Havlin S, Liljeros F, Muchnik L, Stanley HE, Makse HA (2010) Identification of influential spreaders in complex networks. Nat Phys 6(11):888–93
22. Ma LL, Ma C, Zhang HF, Wang BH (2016) Identifying influential spreaders in complex networks based on gravity formula. Phys A Stat Mech Appl 451:205–12
23. Lü L, Zhou T, Zhang QM, Stanley HE (2016) The H-index of a network node and its relation to degree and coreness. Nat Commun 7:10168
24. Sheikhahmadi A, Nematbakhsh MA, Shokrollahi A (2015) Improving detection of influential nodes in complex networks. Phys A Stat Mech Appl 436:833–845
25. Berahmand K, Bouyer A, Samadi N (2019) A new local and multidimensional ranking measure to detect spreaders in social networks. Computing 101(11):1711–33
26. Samadi N, Bouyer A (2019) Identifying influential spreaders based on edge ratio and neighborhood diversity measures in complex networks. Computing 101(8):1147–75
27. Rui X, Yang X, Fan J, Wang Z (2020) A neighbour scale fixed approach for influence maximization in social networks. Computing. 102(2):427–449. https://doi.org/10.1007/s00607-019-00778-5
28. Kumar S, Saini M, Goel M, Panda BS (2021) Modeling information diffusion in online social networks using a modified forest-fire model. J Intell Inf Syst 56(2):355–377
29. Hethcote HW (2000) The mathematics of infectious diseases. SIAM Rev 42:599–653
30. Watts DJ (2002) A simple model of global cascades on random networks. Proc Natl Acad Sci 99(9):5766–71

31. Yamasaki K, Matia K, Buldyrev SV, Fu D, Pammolli F, Riccaboni M, Stanley HE (2006) Preferential attachment and growth dynamics in complex systems. Phys Rev E 74(3):035103
32. Leskovec J, Krevl A, SNAP Datasets (2014) Stanford large network dataset collection, vol 2016, p 49. http://snap.stanford.edu/data
33. Leskovec J, Lang KJ, Dasgupta A, Mahoney MW (2009) Community structure in large networks: natural cluster sizes and the absence of large well-defined clusters. Internet Math 6(1):29–123
34. Leskovec J, Kleinberg J, Faloutsos C (2007) Graph evolution: densification and shrinking diameters. ACM Trans Knowl Discov Data (TKDD) 1(1):2-es
35. Rozemberczki B, Allen C, Sarkar R (2019) Multi-scale attributed node embedding. arXiv preprint arXiv:1909.13021
36. Boguná M, Pastor-Satorras R, Díaz-Guilera A, Arenas A (2004) Models of social networks based on social distance attachment. Phys Rev E 70(5):056122
37. Cho E, Myers SA, Leskovec J (2011) Friendship and mobility: friendship and mobility: user movement in location-based social networks. In: ACM SIGKDD international conference on knowledge discovery and data mining (KDD)
38. McAuley JJ, Leskovec J (2012) Learning to discover social circles in ego networks. In: NIPS, vol 2012, pp 548–556

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

2021 2nd International Conference for Emerging Technology (INCET)
Belgaum, India. May 21-23, 2021

Medical Image Fusion using Lifting based Transform Techniques

Anisha Gupta
Department of Electronics & Communication Engineering
Delhi Technological University
Shahbad Daulatpur, Bawana, Delhi –42
anishagupta772@gmail.com

Priyesh Chaurasia
Department of Electronics & Communication Engineering
Delhi Technological University
Shahbad Daulatpur, Bawana, Delhi –42

Sudipta Majumdar
Department of Electronics & Communication Engineering
Delhi Technological University
Shahbad Daulatpur, Bawana, Delhi –42

2021 2nd International Conference for Emerging Technology (INCET) | 978-1-7281-7029-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/INCET51464.2021.9456179

***Abstract*- Medical image fusion is the means to obtain an enhanced image with better quality and reduced redundancy by fusion of two or more images that may or may not belong to the same modalities. The imaging modalities such as Magnetic Imaging Resonance (MRI), Computed Tomography (CT), Positron Emission Tomography (PET), Single Positron Emission Tomography (SPECT) etc have found their application in diagnosis in the medical domain. The paper presents the use of these imaging modalities pairs (MRI-T1/T2; MRI-PET/ SPECT) to perform medical fusion in transform domain. The paper presents two approaches. 1) It uses the lifting step-based wavelet transform (LWT), a better algorithm that takes significantly less computational time, does distortion-free reconstruction for the Average-Max, Edge-Energy fusion rules. 2) It proposes the new fusion rule to obtain an improved fused image. The fusion rule is applied using the DWT and LWT, and comparisons are made accordingly on the image pairs. Entropy, Peak Signal to Noise Ratio (PSNR), and average gradient are used as performance metrics to assess the image quality.**

I. INTRODUCTION

Image fusion is a technique to derive desirable characteristics from the images used in amalgamation to form an image that contains accurate and more eminent data content. There are several Imaging Modalities like MRI, CT, PET, SPECT and X-rays used in medical diagnosis to provide a diagnosis but in a restricted domain. The X-rays helps in examinations of fractures or related bone-related dislocations and degeneration [1]; CT scan is an advanced version for X-ray that creates a detailed bone image, tendons, and soft-tissue; MRI gives a better diagnosis and perspective on the ailments concerning soft-tissues, joints; and PET and SPECT provide better pathological, metabolic functionality, unlike the other techniques.[1],[2].

Therefore, most of the time CT/MRI images are fused with PET/SPECT to get both the anatomical and metabolic characteristics that increase the applicability and the use of images in medical examination [1].

The Image fusion methodology includes either the transform domain or the spatial domain. The spatial domain techniques work with the pixel of images, manipulating them using logarithmic transform, histogram equalization, or simple averaging techniques [3]. The transform domain works in the frequency domain and is associated with the frequency content of the image [3]. The transform domain techniques convert images into multi-scale or sub-bands. Then, some operator is applied to them to enhance the quality before using the inverse operation.

Various methods over the years have been proposed for Medical image fusion. Bing Huang, Feng Yang, Mengxiao Yin, Xiaoying Mo, Cheng Zhong [4] presents the review for medical fusion in popular domains like non sampled shearlet transform (NSST), DWT, Nonsubsampled Contourlet Transform and (NSCT), NSST domain and different methods including pulse-coupled neural network (PCNN) based fusion, Frei-Chen Operator Medical Image Fusion and deep- learning-based methods. K. KoteswaraRao and K. Veera Swamy [5] used both the transform domain (NSCT, DWT) and the ESOP values in the spatial domain to fuse the images. B. Ashwanth and K. V. Swamy [2] proposed a new fusion algorithm to carry out Medical image fusion in DWT and SWT transform domains highlighting the comparisons between the usage of un-decimated and decimated wavelet transformation. Tan, W., Tiwari, P., Pandey, H.M. *et al.* [6] proposes the new algorithm based on PCNN and energy fusion in NSST domain tested on Alzheimer's disease and related datasets. M. P., T. Singh, R. Nayar and S. Kumar [7] uses CNN architecture for medical image fusion and different fusion methods for denoising the medical images. Jingming Xia, Yi Lu, Ling Tan [8] proposes the new fusion method in the NSST domain that employs the fusion of high-level frequency components using parameter adaptive PCNN and low-level frequency components using convolutional sparse representation (CSR). P.V.Ramaraju, D.Sudharani [9] proposes a new hybrid fusion methods for CT and MRI images in DWT and discrete stationary wavelet transform domain. P. Mathiyalagan [10] proposes Curvlet Transform for medical image fusion that intents to give sharpened image resolution of a fused image. K. Vanitha, D. Satyanarayana and M. N. G. Prasad [11] proposes hybrid l1-l0 decomposition based two scale fusion method and mean, standard deviation methods are employed to evaluate results.

In our paper, to perform the medical image fusion, we use LWT on pre-existing fusion rules like Average-Max and Edge-Fusion to draw comparisons. We also propose a new fusion algorithm that uses both the DWT, a first-generation-based transformation, and the LWT, a second-generation-based wavelet transformation.

DWT has its advantages like the preservation of the spatial features [2],[12], less distortion, and more efficiency in 1-D singularity [2]. The LWT preserves these advantages of DWT and improves the limitations by reducing the error, giving perfect reconstruction, decreasing the computational time and increasing the vanishing moments [13], [14].

II. WAVELET TRANSFORMS

A. Discrete Wavelet Transform (DWT)

The DWT is defined by [12]:

$$W\varphi(j_0,k) = \frac{1}{\sqrt{M}} \sum_x f(x)\varphi_{j_0,k}(x) \quad (1)$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:59:31 UTC from IEEE Xplore. Restrictions apply.

$$W\psi(j,k) = \frac{1}{\sqrt{M}} \sum_k f(x)\psi_{j,k}(x) \quad (2)$$

The coefficient in the first equation is called the approximation coefficient and the coefficient in the second equation is called the detail coefficient.

The 2-D DWT is performing the 1-D wavelet transform two times on a 2-D image. Therefore, at an instance, only one dimension of an image is considered. When DWT is applied on a row, coarse row and detail row sub-bands are created. This process is repeated with the column dimension as well [15].

The sub-band created can either be due to low pass filter or high pass filter. If we produce a sub-band using a low pass filter on a row and a high pass filter on a column, it is denoted as LH. The decomposition of an image gives Low-Low (LL), Low-High (LH), High-Low (HL), High-High (HH) sub-band. Each sub-band gives different kind of information. The bands LH gives the horizontal features, HL gives the vertical features, HH gives the information about the diagonal features and the LL gives the approximation details.

TABLE I. SUB-BAND AND THE INFORMATION IT STORES

$Image_{LL\text{-}BAND}$ Approximation	$Image_{LH\text{-}BAND}$ Horizontal features
$Image_{HL\text{-}BAND}$ Vertical features	$Image_{HH\text{-}BAND}$ Diagonal features

B. Lifting Wavelet Transform (LWT)

The lifting wavelet transform is the second generation wavelet transform, an improvement over the first-generation transform that can perform wavelet mapping from Integer to Integer [13], provide perfect reconstruction, and can be applied on non-separable lattices.

Discrete Wavelet Transform incorporates the calculation of successive averages and differences of a signal. Contrary to this, Lifting Wavelet Transform applies these operations on a part of a signal. This is achieved by three lifting steps Split, Predict and Update [14].

Split - Generally, the signal is split into two disjoint subsets, mainly an odd subset and an even subset.

Predict - Usually, an even subset is used to predict the odd component (subset) via the Prediction operator P. The correlation between the subsets serves as the basis of this prediction. The difference between the actual odd-subset and the subset obtained from the prediction of the even component is computed. This is known as Residual Error(r) or the detail coefficient.[13]

Update - Then Update operator is employed on the r to give coarse approximation data. The data obtained when added to even-subset generates the information equivalent to the information obtained from the signal at low-resolution.

After these three steps, the first level signal upgrade is completed. This is equivalent to the single layer of wavelet-decomposition.[13]

These lifting steps reduce the arithmetic operations by the factor of 2 and significantly decreases the computational time.

III. METHODOLOGY

1. *Average-Max* – In this algorithm, after the decomposition of an image, the average operation is applied on the LL band and maximum operation on the other bands before the inverse operation is applied to create a fused image.[16]
2. *Max-Max* - In this algorithm, a max operation is applied on all the frequency bands, the low and high-frequency bands to construct a fused image.
3. *Edge-Energy* – This algorithm calculates the edge information using the Sobel operator in the LL band, that is low-level sub-band and using the energy fusion rule in the remaining LH, HH, and HL sub-bands to create a fused image [2]. The energy fusion rule for the band is:

$$E(j,k) = \sum_{j-1}^{j+1} \sum_{k-1}^{k+1} (LH/HL/HH(j,k)) \quad (3)$$

IV. PROPOSED METHOD

We use different image pairs to perform the image-fusion. In this paper, we have used the MRI_T1-MRI_T2, MRI-PET, and MRI-SPECT image modalities to evaluate the accuracy of both the approaches.

1) LWT using the Average-Max and Edge-Energy fusion rule.

2) LWT and DWT using the proposed fusion rule.

We preprocess the image by converting it into grayscale and making the size of both the images equal.

We have done the decomposition and reconstruction process using 'sym5' wavelet.

A. LWT on Edge-Energy and Avergae-Max

This paper uses the two-dimensional lifting decomposition to 'sym5' wavelet to construct the fused image. We use two fusion rules, Average-Max and Edge-Energy Fusion rules, and then we make comparisons with DWT using these same rules [3].

B. Proposed Hybrid Approach using DWT and LWT

Step 1: We first perform the Edge-Edge fusion on Input Image A and Input Image B to get the intermediary fused image. The edge-edge fusion rule calculates the edge information on all the sub-bands. In this particular edge-edge, we calculate edge information using the Sobel Operator.

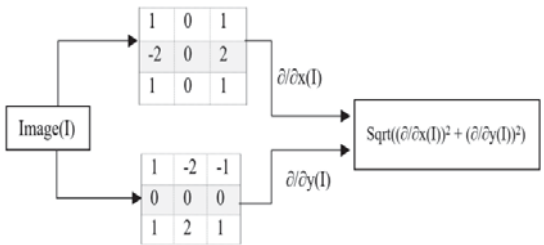


Fig. 1. Sobel Edge Detection

Step 2: We fuse the fused image obtained in step1 with the input image A to create a second fused image using the Average-Maximum fusion rule.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:59:31 UTC from IEEE Xplore. Restrictions apply.

Step 3: We fuse the fused image obtained in step1 with the input image B to create a third fused image using the Average-Maximum fusion rule.

Step 4: We create a final image using the images obtained in Step-2 and Step-3 using the Maximum-Maximum rule.

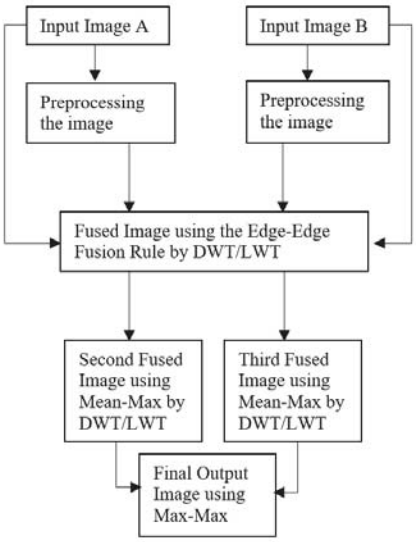


Fig. 2. Proposed Approach

V. RESULTS

A. Performance Metrics

1) Entropy – It is used to determine the information content in an image. Therefore, the greater entropy is desirable.

2) PSNR - It is defined as the ratio between the maximum signal power of the reference signal and the corrupting noise. It is expressed in the decibels (dB). If A be the reference signal/image and B be the fused/output image. The PSNR is given by

$10\log_{10}(N^2/\text{mean-square-error(A,B)})$.

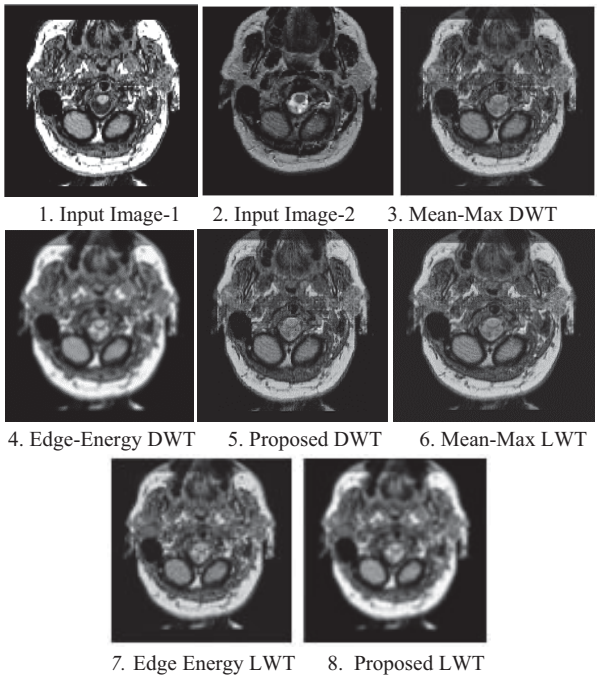
If error is less, PSNR value is high.

3) Average Gradient - We characterize the clarity and the quality of the image using the average gradient. Therefore, the greater the gradient, the better the quality of the fused image.[17]

$$g = \frac{1}{(M-1)(N-1)} \sum_{i=1}^{(M-1)(N-1)} \sqrt{\frac{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}{2}}$$

B. OUTPUT – Code and Simulations are done through MATLAB.

SET-1 – MRI_T1-MRI_T2



1. Input Image-1 2. Input Image-2 3. Mean-Max DWT

4. Edge-Energy DWT 5. Proposed DWT 6. Mean-Max LWT

7. Edge Energy LWT 8. Proposed LWT

Fig. 3. Output of Set-1

1)1st Input Image- MRI – T1, 2) 2nd Input Image – MRI – T2 3)Output with Mean-Max DWT, 4) Output Image with Edge-Energy DWT, 5) Proposed approach – DWT, 6) Output image with Mean-Max LWT, 7) Output with Edge-Energy LWT, 8) Output image with proposed approach - LWT

TABLE : II COMPARISON OF PREVIOUS APPROACHES WITH OUR NEW APPROACH USING DWT AND LWT.

		DWT			LWT		
		Mean-Max	Edge-Energy	Proposed	Mean-Max	Edge-Energy (Approach-1)	Proposed (Approach-2)
Entropy	Set-1	1.6558	1.7925	1.7363	1.7233	1.8771	1.9448
	Set-2	2.0153	2.0987	2.1165	2.0318	2.2205	2.1663
	Set-3	1.4686	1.4325	1.5119	1.5308	1.5204	1.6054
	Set-4	1.5405	1.6679	1.6681	1.5770	1.7733	1.8001
PSNR	Set-1	18.1520 18.2035	11.6834 12.9241	13.7846 13.7846	18.2036 18.1515	11.4756 11.4602	11.9433 18.7769
	Set-2	20.4311 20.4311	11.5390 14.6748	18.0022 19.0945	20.4309 20.4313	11.5396 14.6704	18.0018 19.0951
	Set-3	15.3647 15.3827	11.4913 8.5450	15.5894 13.9051	15.3831 15.3643	11.3644 8.4624	15.6091 13.8919
	Set-4	17.3891 17.4018	6.9787 9.7437	13.0504 5.2784	17.4015 17.3894	6.9912 9.7726	11.3609 15.9802
Average Gradient	Set-1	56.9549	39.9167	42.6392	57.3730	43.5732	43.1929
	Set-2	27.1991	21.4137	23.2190	27.2273	21.7885	26.2621
	Set-3	37.6370	25.6572	31.0977	37.9434	28.7622	31.3981
	Set-4	46.4641	37.0165	36.8755	46.6557	36.3363	37.1184

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:59:31 UTC from IEEE Xplore. Restrictions apply.

SET-2 MRI-SPECT

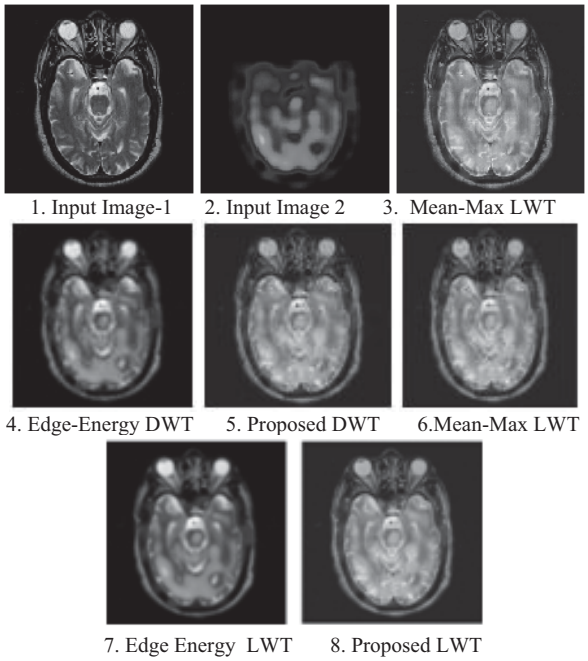


Fig. 4. Output of Set-2
1)1st Input Image- MRI , 2) 2nd Input Image – SPECT 3)Output with Mean-Max DWT, 4) Output Image with Edge-Energy DWT, 5) Proposed approach – DWT, 6) Output image with Mean-Max LWT, 7) Output with Edge-Energy LWT, 8) Output image with proposed approach - LWT

SET - 3 MRI-PET

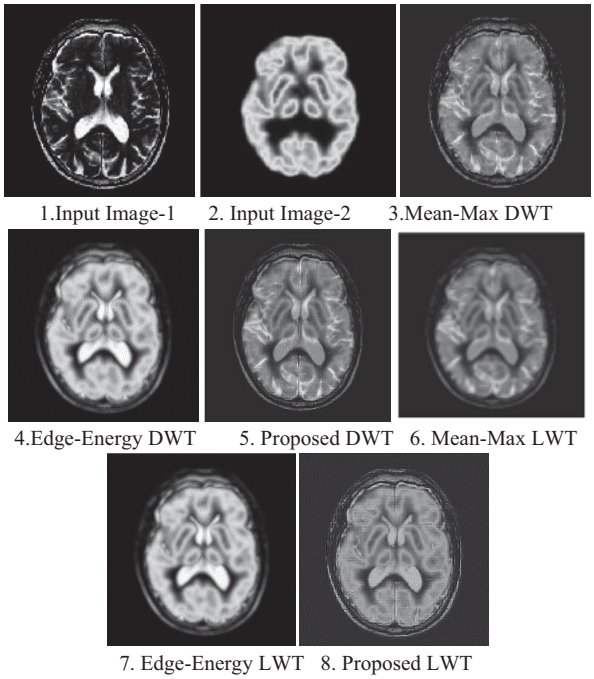


Fig. 5. Output of Set-3
1)1st Input Image- MRI , 2) 2nd Input Image – PET 3)Output with Mean-Max DWT, 4) Output Image with Edge-Energy DWT, 5) Proposed approach – DWT, 6) Output image with Mean-Max LWT, 7) Output with Edge-Energy LWT, 8) Output image with proposed approach - LWT

SET-4 MRI_T1-MRI_T2

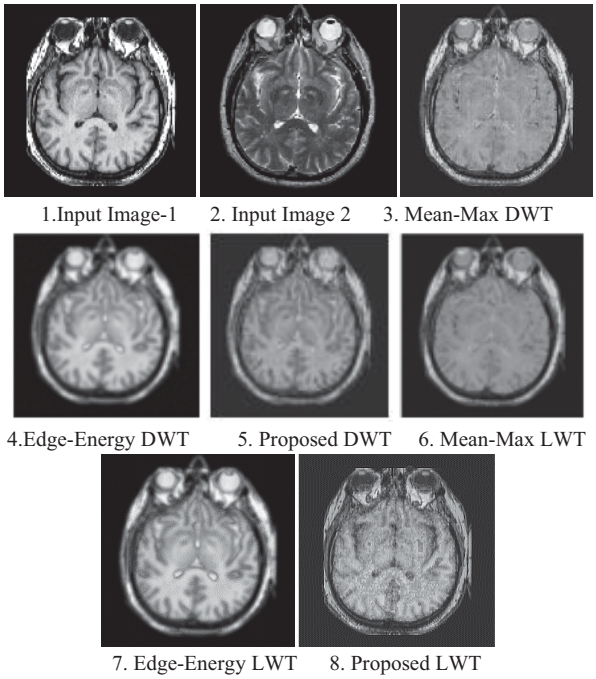


Fig. 6. Output of Set-4
1)1st Input Image- MRI-T1, 2) 2nd Input Image – MRI-T2 3)Output with Mean-Max DWT, 4) Output Image with Edge-Energy DWT, 5) Proposed approach – DWT, 6) Output image with Mean-Max LWT, 7) Output with Edge-Energy LWT, 8) Output image with proposed approach – LWT

VI. CONCLUSION

In terms of entropy, LWT performs better on all the image pairs than the DWT counterparts. Our proposed approach for LWT-Edge Energy works better than existing DWT-Edge Energy. Our proposed fusion algorithm (Approach-2) for LWT works best among all the scenarios.

Similarly, the average gradient calculated using the LWT performs better than the DWT counterparts in all the image-pairs.

The table shows the two values of PSNR, PSNR of fused image with Input Image-1 and PSNR of fused image with Input Image-2.

REFERENCES

[1] https://www.envrad.com/difference-between-x-ray-ct-scan-and-mri/

[2] B. Ashwanth and K. V. Swamy, "Medical Image Fusion using Transform Techniques," 2020 5th International Conference on Devices, Circuits and Systems (ICDCS), Coimbatore, India, 2020, pp. 303-306, doi: 10.1109/ICDCS48716.2020.243604.

[3] Snehal O.Mundhada, V. K.Shandilya, " Spatial and Transformation Domain Techniques for Image Enhancement ", International Journal of Engineering Science and Innovative Technology (IJESIT) Volume 1, Issue 2, November 2012, ISSN: 2319-5967 ISO 9001:2008 Certified.

[4] Bing Huang, Feng Yang, Mengxiao Yin, Xiaoying Mo, Cheng Zhong, "A Review of Multimodal Medical Image Fusion Techniques", Computational and Mathematical Methods in Medicine, vol. 2020, Article ID 8279342, 16 pages, 2020. https://doi.org/10.1155/2020/8279342

[5] K. KoteswaraRao and K. Veera Swamy, " Multimodal Medical Image Fusion using NSCT and DWT Fusion Frame Work", Volume-9 Issue2, December 2019,pp3644-3645.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:59:31 UTC from IEEE Xplore. Restrictions apply.

[6] Tan, W., Tiwari, P., Pandey, H.M. et al. Multimodal medical image fusion algorithm in the era of big data. Neural Comput & Applic (2020). https://doi.org/10.1007/s00521-020-05173-2

[7] M. P., T. Singh, R. Nayar and S. Kumar, "Multi Modal Medical Image Fusion using Convolution Neural Network," 2019 Third International Conference on Inventive Systems and Control (ICISC), Coimbatore, India, 2019, pp. 351-357, doi: 10.1109/ICISC44355.2019.9036373

[8] Jingming Xia, Yi Lu, Ling Tan, "Research of Multimodal Medical Image Fusion Based on Parameter-Adaptive Pulse-Coupled Neural Network and Convolutional Sparse Representation", Computational and Mathematical Methods in Medicine, vol. 2020, Article ID 3290136, 13 pages, 2020. https://doi.org/10.1155/2020/3290136

[9] P.V.Ramaraju , D.Sudharani, "Medical Image Fusion using Discrete Wavelet Transform and Proposed Hybrid technique", International Journal for Research in Applied Science & Engineering Technology (IJRASET), ISSN: 2321-9653, Volume 6 Issue V, doi: http://doi.org/10.22214/ijraset.2018.6139.

[10] P. Mathiyalagan, "Multi-Modal Medical Image Fusion Using Curvelet Algorithm," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Bangalore, India, 2018, pp. 2453-2458, doi: 10.1109/ICACCI.2018.8554759.

[11] K. Vanitha, D. Satyanarayana and M. N. G. Prasad, "Multimodal Medical Image Fusion Based on Hybrid L1- L0 Layer Decomposition Method," 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 2019, pp. 1-5, doi: 10.1109/ICCCNT45670.2019.8944896.

[12] (2008) Discrete Wavelet Transform (DWT). In: Furht B. (eds) Encyclopedia of Multimedia. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-78414-4_305.

[13] Yu Guang and Zhao Xue Jun, "The study of the second generation of wavelet transform," World Automation Congress 2012, Puerto Vallarta, Mexico, 2012, pp. 1-4.

[14] La Cour–Harbo, A., & Jensen, A. (2009). Wavelets and the Lifting Scheme. Encyclopedia of Complexity and Systems Science, 10007–10031. doi:10.1007/978-0-387-30440-3_588

[15] Mistry, Dr. (2013). DISCRETE WAVELET TRANSFORM USING MATLAB. International journal of Computer Engineering & Technology (IJCET). 4. 252-259.

[16] Meda Balachandra Mule and Padmavathi N.B,"Basic Medical Image Fusion Methods", 2015,pp1046-1048

[17] El-Hoseny, H. M., El.Rabaie, E.-S. M., Elrahman, W. A., & El-Samie, F. E. A. (2017). Medical image fusion techniques based on combined discrete transform domains. 2017 34th National Radio Science Conference (NRSC). doi:10.1109/nrsc.2017.7893518

[18] Jagalingam Pa and Arkal Vittal Hegdeb, " A Review of Quality Metrics for Fused Image",2015.135-137

[19] TianLan, ZheXiao, YiLi,Yi Ding, and Zhiguang, |"Medical image fusion techniques based on combined discrete transform domains",2016,pp788-881

[20] Swathi.P.S, Sheethal.M.S, Vince Paul, "Survey on Multimodal Medical Image Fusion Techniques", IJCSET, vol 6, Issue 1, 33-39, 2016.

[21] Z. Wang and A. Bovik, "A universal Image Quality Index," IEEE Signal Processing Letters, vol. 9, no. 3, pp. 81-84, 2002.

[22] C. Yang, J. Zhang, X. Wang, X. Liu, "A novel Similarity based Quality Metric for Image Fusion," Information Fusion, vol. 9, no. 2, pp. 156-160, 2008.

[23] Ijcaonline.org

[24] www.rroioj.com

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:59:31 UTC from IEEE Xplore. Restrictions apply.

Communicative & Integrative BIOLOGY

Volume 9 • Issue 4 • 2016



Communicative & Integrative Biology

Taylor & Francis
Taylor & Francis Group

Taylor & Francis
ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/kcib20

Microbial Amelioration of Salinity Stress in HD 2967 Wheat Cultivar by Up-Regulating Antioxidant Defense

Madhulika Singh & Neha Tiwari

To cite this article: Madhulika Singh & Neha Tiwari (2021): Microbial Amelioration of Salinity Stress in HD 2967 Wheat Cultivar by Up-Regulating Antioxidant Defense, Communicative & Integrative Biology, DOI: 10.1080/19420889.2021.1937839

To link to this article: https://doi.org/10.1080/19420889.2021.1937839

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Accepted author version posted online: 01 Jun 2021.

Submit your article to this journal

Article views: 25

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=kcib20

Check for updates

Publisher: Taylor & Francis & Informa UK Limited, trading as Taylor & Francis Group

Journal: *Communicative & Integrative Biology*

DOI: 10.1080/19420889.2021.1937839

Microbial Amelioration of Salinity Stress in HD 2967 Wheat Cultivar by Up-Regulating Antioxidant Defense

Madhulika Singh[a*] and Neha Tiwari[b]

[a]Department of Botany, SSN College, University of Delhi, Delhi-110036, India; [b]Department of Biotechnology, Delhi Technological University, Delhi-110042, India

***CONTACT:** M. Singh, Department of Botany, SSN College, University of Delhi, Delhi-110036, India; Email: madhulikasingh@ss.du.ac.in

ACCEPTED MANUSCRIPT

ABSTRACT

An experiment was conducted to investigate the potential of *Piriformospora indica* and plant growth-promoting bacteria (PGPB) to ameliorate salinity stress in HD 2967 wheat cultivar. Plants were treated with four different levels of salinity viz. 0, 50, 100, and 200 mM NaCl (electrical conductivity value 0.01, 5.84, 11.50 and 21.4 mS cm^{-1}, respectively) under greenhouse conditions, using a completely randomized design experiment. Plants inoculated with PGPB and *P. indica* showed decrease in lipid peroxidation, relative membrane permeability and lipoxygenase enzyme (LOX) activity as compared to uninoculated plants. The result of this study showed that PGPB and *P. indica* inoculated HD 2967 wheat plants accumulated higher content of proline, α-tocopherol and carotenoid as compared to uninoculated plants. The HD 2967 wheat plants either inoculated with PGPB or *P. indica* showed significantly higher activities of antioxidant enzymes viz. superoxide dismutase, catalase and ascorbate peroxidase than that of the uninoculated plants. Moreover, PGPB inoculated plants showed greater activity of antioxidant enzymes than the plants inoculated with *P. indica*. Salinity stress tolerance was more pronounced in the PGPB inoculated than *P. indica* inoculated plants. This study revealed the potentiality of PGPB and *P. indica* as bio-

ameliorator under salinity stress, and suggests that this plant microbial association could be a promising biotechnological tool to combat the deleterious effects of salinity stress.

KEYWORDS: Salinity, antioxidants, plant growth promoting bacteria, *P. indica*, abiotic stress

Introduction

As per an estimation of World Population Prospects (2020), the human population is currently increasing at the rate of 1.05% per year, and it is expected to be around 7.8 billion in 2021 Agriculture sector is facing a major problem to meet the food requirement of growing population. However, plant development and yield has been severely influenced by several abiotic stresses like salinity, drought, temperature, and heavy metals [1]. Among various abiotic stresses, salinity is one of major deleterious factors, which limits crop productivity, in various parts of world [2]. About 20% of cultivated land area is affected by salinity globally [3]. According to an estimate, about 50% of cultivated land will be affected by salinity by 2050, if salinization process continues at a current pace [4]. Agricultural land is affected by salinization primarily due to the excessive accumulation of soluble salts in the soil, particularly due to accumulation of NaCl salt [1].

Salinity stress often induces overproduction of reactive oxygen species (ROS) including hydroxyl radical ($OH^{\bullet}$), single oxygen ($^1O_2$), superoxide anion ($O_2^{\bullet-}$) and hydrogen peroxide (H_2O_2) [5]. ROS formation causes oxidative damage to proteins, lipids and photosynthetic pigments [6]. ROS deteriorates the structure of cell membrane by oxidation of polysaturated fatty acids of the lipid layer, altering their permeability. However, plants have an excellent antioxidant system to ameliorate ROS induced oxidative stress via enzymatic and non-enzymatic antioxidants production [7,8]. Several studies have revealed that the enzymatic antioxidant system (viz. ascorbate peroxidase (APX), catalase (CAT), superoxide dismutase (SOD), and the non-enzymatic antioxidants like as carotenoids, proline, α-tocopherol and ascorbic acid could combat oxidative stress and thereby facilitating plants to better adopt under saline conditions [7,8].

The level of lipid peroxidation largely increased under saline conditions [9], and has an influence on the plant membrane properties, protein degradation and the capacity of ion transport, and eventually triggers the process of plant cell death. However, ROS production

ACCEPTED MANUSCRIPT

under salt stress further initiates higher production of lipid peroxidation [10]. Excessive ROS production and peroxidation of lipids alters the activity of lipoxygenase enzymes [11], therefore enhance the conversion of lipids into hydroperoxyl fatty acids [12]. Salinity induced enhanced LOX activity subsequently led to cell death [11]. In turn, proline is an important scavenger of ROS, providing protection against oxidative damage under salinity stress [13] and APX imparts a defensive role against ROS [14]. Moreover, α-tocopherol scavenge free radicals responsible for oxidation of polyunsaturated fatty acids hence, protect cell membrane from lipid peroxidation and maintain membrane integrity in host plants [15]. Ascorbate is a major antioxidant molecule playing a crucial role in detoxification of ROS [16], protection of proteins and lipids during salinity stress [17]. Agami [18] reported that ascorbate application in barley improved plant growth, relative water content, proline, photosynthetic pigments, enzymatic antioxidants under salt stress. Under salinity stress polyphenols accumulates in various plant tissues and scavenges free radicals [19]. Polyphenols regulate plant growth, pigment synthesis, scavenge ROS which help the plant to mitigate abiotic stresses viz. salinity, drought, temperature, [20]. Carotenoids are the harvester of light during photosynthesis. Moreover, they act as chloroplast membrane stabilizers hence reducing ion leakage and peroxidation of lipids [21,22]. The enzyme SOD catalyses the conversion of $O_2^{\cdot -}$ to H_2O_2 [8] thus, plays a significant role in protecting plant from oxidative damage [23]. Catalase reduces the level of ROS by catalysing the breakdown of hydrogen peroxide into water and oxygen [24].

In the recent years, several salt tolerant crops have been developed by conventional breeding programme and recombinant DNA technologies, however, have been proven to be time consuming and cost intensive. In light of upcoming challenges, there is a need to develop alternative technologies for sustainable agriculture production, such as the use of soil microbes [25]. Several reports have revealed that the soil inoculation with beneficial soil microorganisms increases the production of antioxidant enzymes such as, SOD, CAT, and peroxidase (POD) which scavenge excess free radicals produced during salt stress [26] thus, elevate salt tolerance in crop plants and increase their growth. Amongst a wide array of such microorganisms, *P. indica* inoculation has been found to reduce salinity stress in rice plants [27] and drought stress in strawberry plants [28]. Further, *P. indica* colonization increases the activity of antioxidant defense enzymes in the plants exposed to salinity [29]. PGPB (*Bacillus cereus*) confer salinity tolerance in *Vigna radiata* by enhancing the activity of SOD, CAT and

ACCEPTED MANUSCRIPT

peroxidase which belongs to antioxidant defense system [30]. *Azotobacter chrococcum* alleviated salt stress in maize by increased accumulation of proline and polyphenols [31]. *P. indica* inoculation also enhance the accumulation of non-enzymatic antioxidants such as carotenoids and proline in rice plants [27] while grow under saline conditions. Recently, Kasotia et al. [32] reported that *Pseudomonas koreensis* colonization in soyabean plants stimulated the SOD and CAT enzyme activities under abiotic stress. The higher proline content has also been recorded in the plants colonized with arbuscular mycorrhizal fungi [33]. The α-tocopherol a vitamin E compound located in the membrane of chloroplast and thylakoid and also in the plastoglobuli. This antioxidant deactivates ROS generated during photosynthesis (mainly $^1O_2$ and $OH^{\bullet}$), and lowers rate of MDA production by scavenging lipid peroxyl radicals [34]. The α-tocopherol level changes in response to abiotic stress. High α-tocopherol content in microbial inoculated plants lowers lipid peroxidation [26] as the α-tocopherol involves in scavenging lipid peroxyl radicals. PGPB enhanced salinity tolerance in plants by increasing the production of non- enzymatic antioxidants such as α-tocopherol, ascorbate, carotenoids and polyphenols [35]. *P. indica* colonization in maize plant decreased the level of MDA production by stimulating antioxidant enzymes [36].

HD 2967 (Pusa Sindhu Ganga), a high yielding double-dwarf wheat variety, is one among the foremost necessary cereal crops in India. The plant growth and productivity are adversely affected by salinity. Till now, no systematic research has explored the effects of using *Azotobacter chroococcum, Enterobacter asburiae and Lactococcus lactis* in combination, on salt-stressed HD 2967 wheat cultivar. As per the literature survey there is no report on *P. indica* induced salinity tolerance in the HD 2967 plant. Therefore, the present study was carried out with an aim to study the effect of *P. indica* and PGPB inoculation in HD 2967 wheat tolerance to salinity induced by NaCl treatment. It is hypothesized that *P. indica* and PGPB protect wheat plant from damaging effects of salt stress by enhancing antioxidant enzyme activities and antioxidant molecule accumulation.

2. Materials and Methods

2.1. Experimental set up, Growth Conditions, and Soil

A pot experiment was conducted under semi-control conditions (Temperature: 24-29 °C; relative humidity: 60 ± 10%) in the botanical garden of Swami Shraddhanand College, University of Delhi, Delhi, India. Humidity was controlled by adequate spacing between plants, regular

weeding, well drained greenhouse floor, opening greenhouse door at night, installation of fans in greenhouse. The temperature was controlled by covering the green house by plastic sheets and installation of LED lights for uniform light distribution in greenhouse. Controlled humidity level discouraged the growth of pests. The soil used in the experiment was autoclaved to avoid contamination of soil.

The soil was collected from the fields of Narela Village, New Delhi, India after consulting from farmers. Six points in the fields were selected. Removed the vegetation and surface litter and 1 cm layer of soil. Soil was collected from a depth of 1-20 cm and placed in sacks and brought to green house. Soil was mixed and kept on a raised surface lined with plastic sheets. Soil was lightly pounded, sieved and air dried for two weeks. Soil was autoclaved at 121°C and 15 psi. The mechanical analysis of soil was done by sieve method and it was found that the percentage of very coarse particles was 2.84%, coarse sand 57.9%, fine sand 16.6% and silt and clay 22.66%. The physico-chemical properties of the soil were determined by the department of soil sciences and agricultural chemistry, IARI, Delhi, India. The physicochemical properties of experimental soil are given in the **Table 1**.

Before conducting an experiment, the soil was sterilized by autoclaving it for 30 minutes at 121°C and 15 psi to kill the prevailing microbes, if any. *Azotobacter chroococcum* is nitrogen solubilizer, *Enterobacter asburiae* is potassium solubilizer and *Lactococcus lactis* is phosphorus solubilizer. The experiment was laid out in a randomized complete block design with 4 factors: i) non-microbial control ii) inoculation with *P. indica* iii) inoculation with PGPB *(Azotobacter chroococcum, Enterobacter asburiae and Lactococcus lactis)* iv) four salinity levels (0, 50, 100 and 200 mM NaCl). Hence, there were 12 treatments (3×4) and each treatment had three replicates.

Table 1. Physicochemical properties of the experimental soil

Parameter	Properties /Content
Soil type	Sandy loam (coarse particles: 2.84%, coarse sand: 57.9%, fine sand: 16.6% and silt and clay: 22.66%)
pH	7.5
Electrical conductivity	0.19 dS m^{-1}
Organic matter	1.1 %

Available N	196 mg kg^{-1}
Available P	51.5 mg kg^{-1}
Available K^{+}	303 mg kg^{-1}
Mg^{2+}	245 mg kg^{-1}
Zn^{2+}	7 mg kg^{-1}
Fe^{3+}	12.5 mg kg^{-1}
Cu^{2+}	4.04 mg kg^{-1}
Mn^{2+}	7.40 mg kg^{-1}.
Na^{+}	320 mg kg^{-1}.

NUSCRIPT

2.2. ***Plant Material and Inoculation Treatment***

Seeds of HD-2967 wheat cultivar used in the study, which were collected as a gift from the Department of Genetics, Indian Agricultural Research Institute (IARI), New Delhi, India. The seeds were surface sterilized by treating the seeds with 4% sodium hypochlorite for 1 minute and then washed three times with Milli Q water to remove any trace of chemical that could interfere with seed germination.

The sterilized seeds were directly sown in the earthenware pots containing 4 kg of autoclaved soil at a depth of 3 cm. Plants were watered twice a week. PGPB inoculum was obtained from the Department of Microbiology, IARI, New Delhi, India, and *P. indica* inoculum was procured from the Institute of Microbial Technology, Amity University, Noida, India. The microbes were added in the soil at the time of seed sowing at a depth of 3 cm. Same amount of sand was added in non-inoculated soil. Added recommended dose of fertilizer for pot experiment viz: nitrogen (N) as urea: 187.5 kg ha^{-1}, diammonium phosphate (P): 75 kg ha^{-1}, and muriate of potash as (K): 75 kg ha^{-1}. One-third dose of nitrogen and complete doses of P and K were added in the soil at the beginning of experiment. The rest amount of nitrogen was equally divided and added after 21 days and 42 days of seed sowing, respectively.

2.3. Saline Stress Treatment

Plants were watered with autoclaved tap water so that adequate moisture is maintained in the earthenware pots. Plants were allowed to grow for 20 days prior to the salt treatment to allow microbes to colonize. The plants were treated with four different levels of salinity stress (0, 50, 100 and 200 mM NaCl) once a week with 50 ml of each NaCl solution per pot for 60 days and a total of 400 ml of NaCl solution was added to each pot during the entire experiment. The electrical conductivity (EC) of soil was measured using conductivity meter (HACH analyzer, HQ440d multi) and it increased to 0.19, 5.84, 11.50, 21.4 mS cm^{-1} in the 0, 50, 100 and 200 mM NaCl salinity levels, respectively. The experiment was carried out under green house condition (Light intensity: 1000 lux; Temperature: 24-29 °C; relative humidity: 60 ± 10). Plants were harvested by uprooting the entire plant manually after 80 days of sowing at stem extension second node visible seventh phenological stage of wheat development for analysis (**Figure 1)**.

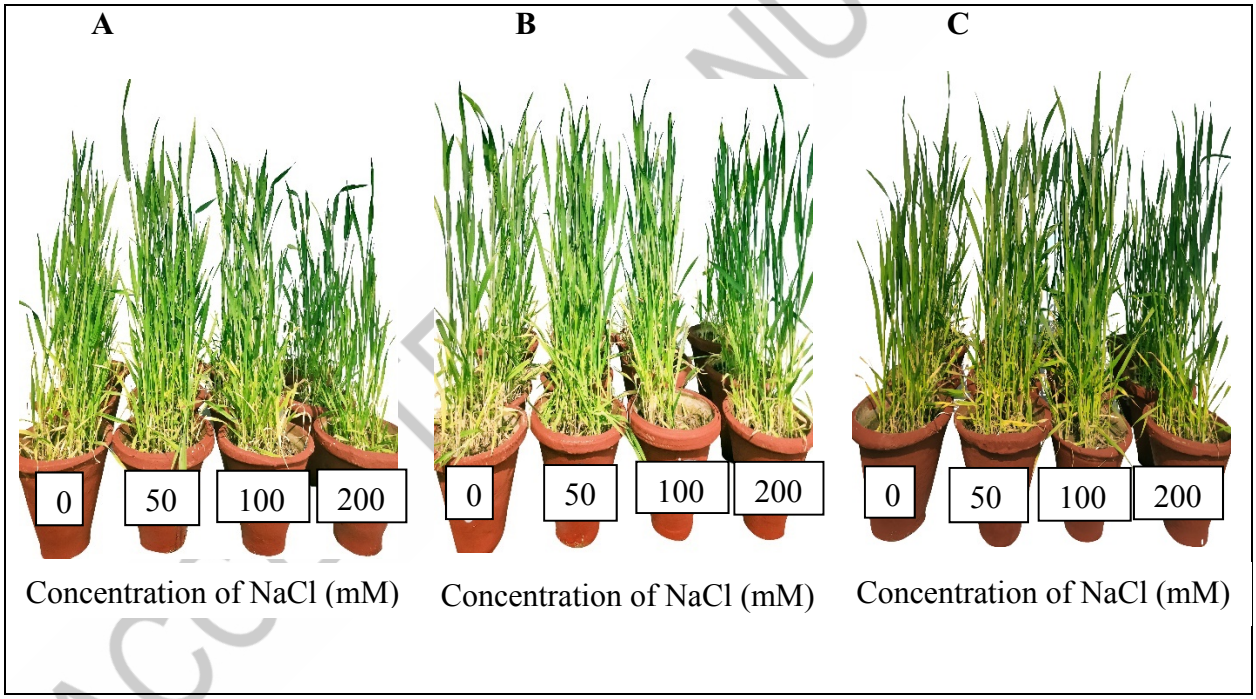


Figure 1 Experimental set-up showing (**A**) un-inoculated wheat plants (control); (**B**) *P. indica* inoculated wheat plants; and (**C**) PGPB inoculated wheat plants after 80 days of sowing.

2.4. *Root and Shoot Biomass Estimation*

ACCEPTED MANUSCRIPT

Plants were harvested after 80 days of seed germination. Root and shoot samples were oven dried two days at 80°C to obtain constant weight and then root and shoot biomass was determined based on weights of the dried samples.

2.5. *Oxidative Damage Estimation*

The membrane relative permeability of leaves was determined following the protocol of Zwiazek and Blake [37]. Second leaf from top of the plants were rinsed with Milli Q water to eliminate electrolytes sticked to the surface of leaves. Leaf discs of 5 mm diameter were cut out. The leaf discs were submerged in test tubes containing 25 ml of Milli Q water. After incubation for one-hour, electrical conductivity of the solution was measured using conductivity meter (HACH analyzer, HQ440d multi). Thereafter, the leaf discs were boiled to allow lysis all cells. The solutions were cooled and conductivity was measured [37].

$$\text{Relative permeability} = \frac{\text{EC of the solution before heating}}{\text{EC of the solution after heating}} \times 100$$

The lipid peroxidation level in leaves was estimated in terms of malondialdehyde content (MDA), following the methods of Heath and Packer [37]. Leaf tissue (0.5 g) was crushed in 0.1% trichloroacetic acid (TCA). The homogenate was centrifuged at 12000 rpm for 15 min at (Eppendorf, Centrifuge 5810 R). To the supernatant liquid 0.5% thibarbituric acid (TBA) in 20% TCA was added. The mixture was heated and then cooled. The mixture was centrifuged for 10 min [38]. The absorbance reading of the supernatant was noted at 532 nm (Eppendorf UV-vis Spectrophotometer, BioSpectrometer basic model). The calculation of the MDA content was carried out using its extinction coefficient of 155 mM^{-1} cm^{-1} and expressed as n mole g^{-1} fw [38].

Lipoxygenase (LOX) activity was measured according to the protocol given by Doderer et al. [39]. Leaf tissue (500 mg) was crushed in phosphate buffer containing 0.5 mM EDTA. The homogenate was centrifuged at 12,000 rpm for 15 min. The enzyme activity was determined by adding 50 µL extract to 2.95 mL substrate, which was prepared by adding 70 µL of linoleic acid to 10 mL milli Q water containing 100 µL tween 20, final volume adjusted to 200 mL by using 0.1 M phosphate buffer solution [39]. Absorbance reading was noted at 234 nm and the activity of enzyme was expressed in µmol $min^{-1}g^{-1}$ fw.

ACCEPTED MANUSCRIPT

2.6. ***Antioxidant Molecules Estimation***

Leaf tissue (200 mg) was homogenized in 3% aqueous sulfosalicylic acid to determine the content of proline [40]. The homogenate was centrifuged at 12000 rpm for 10 min. The essay mixture contained supernatant (2000 µl), acid ninhydrin, and anhydrous acetic acid. The mixture was boiled for 1 hour and cooled immediately. Chromophore was extracted by adding toluene and absorbance was determined spectrophotometrically at the wavelength of 520 nm [40]. The proline content was determined from a standard curve and expressed in µmole g^{-1}.

Leaf tissue (100 mg) was crushed in liquid nitrogen and homogenized in 0.1 N H_2SO_4 to determine the content of α-tocopherol [41]. The homogenate was centrifuged at 10,000 rpm for 10 min at room temperature. An equal amount of ethanol was added to the supernatant and centrifuged for 5 min. To the mixture, equal volume of xylene was added and centrifuged at for 5 min. To 1 mL of upper phase, 0.12% (w/v) $FeCl_3$ was added and incubated at room temperature [41]. The absorbance was observed at 520 nm wavelength. The content of α-tocopherol was expressed as µg g^{-1}fw.

The content of carotenoids in plant leaves was estimated according to Hiscox and Israelstam [42]. Leaves (100 mg) were chopped into small pieces and it was dipped in 36 test tubes containing 7ml dimethyl sulfoxide. The tubes were kept at 65$^{\circ C}$ for 4 h and final volume was made to 10 ml with dimethyl sulfoxide. The absorbance reading was recorded at 480 and 510 nm and the content of carotenoid was calculated as per the formula given by Arnon [43] and expressed as mg g^{-1}fw.

2.7. Analysis of ***Antioxidant Enzymes***

Leaf tissue (500 mg) was crushed in 0.1 M phosphate buffer (pH 7.5) solution containing 0.5 mM EDTA. The homogenate was centrifuged at 12,000 rpm for 20 min and the extract was used to measure enzyme activity.

SOD activity was determined following the protocol of Dhindsa et al. [44]. The reaction mixture contained 1.5 M sodium carbonate, 2 mM methionine, 2.25 mM Nitro-blue tetrazolium, 3 mM EDTA, 100 mM potassium phosphate buffer, Milli Q water and 50 µL of enzyme. The reaction was started by adding 60 µM riboflavin and the test tubes were placed

ACCEPTED MANUSCRIPT

under 15 W florescent lamps for 15 min. The absorbance was read at 560 nm after the reaction was stopped [44]. The unit of enzyme activity was considered as the amount of enzyme in leaf tissue which decreased the absorbance to fifty percent as compared to the test tubes not having enzymes. The enzyme activity was expressed as unit min^{-1} g^{-1}fw.

The CAT activity was determined following the protocol of Teranishi et al [45]. The reaction mixture contains 6 mM H_2O_2, 0.1 M phosphate buffer (pH 7.0) and enzyme extract. The reaction was terminated using titanium reagent. The mixture was centrifuged at 10,000 rpm for 10 min and absorbance was noted at 415 nm [45]. The enzyme activity was expressed in unit min^{-1} g^{-1}fw.

The APX activity was measured according to Nakano and Asada protocol [46]. To three mL of reaction mixture having potassium phosphate buffer (100 mM), ascorbic acid (3 mM), EDTA (3 mM), and H_2O_2 (6 mM), the enzyme extract was added for estimation of APX activity [46]. The change in absorbance was read at 290 nm and expressed as unit min^{-1} g^{-1} fw.

2.8. *Statistical Analysis*

The data were analysed using SPSS 21 statistical programme (IBM SPSS Statistics 21) by one way ANOVA with NaCl treatment, microbial inoculation and interactions among them as a source of variation. Comparison of the means were determined by post hoc Duncan's test ($p < 0.05$).

3. Results

3.1. *Shoot and Root Biomass*

Microbial inoculation of HD 2967 wheat variety showed higher plant growth at all levels of salinity as compared to the uninoculated plants. However, NaCl treatment (200 mM) lowered the shoot by 3.6 % compared to plants without NaCl treatments. The root biomass remained unchanged at 0 mM NaCl treatment as well as 200 mM NaCl treatment in uninoculated control plants, and decreases proportionally in case of *P. indica* inoculation. Root biomass appeared same at 50 and 100 mM NaCl, and a very diminutive difference was detected at 200 mM NaCl treatment in case of PGPB inoculation. On comparing the potential of PGPB and *P. indica* it was observed that PGPB inoculation exhibited a higher shoot and root biomass than *P. indica* colonized plants. At 200 mM NaCl salinity level, PGPB inoculation increased

the shoot and root biomass by 48 % and 51.25 %, respectively while *P. indica* inoculation increased it by a 24.4 % and 30.89 % respectively, compared to un-inoculated plants (**Figure 2 A-B**). Biomass of shoot increased with increasing levels of salinity (0, 50 and 100 mM NaCl) with microbial inoculation; however, biomass decreased when wheat plants were supplied with 200 mM NaCl solution in both microbial inoculated and uninoculated plants. The dry weight was significantly greater in PGPB and *P. indica* inoculated under nonsaline conditions in comparison to uninoculated plants.

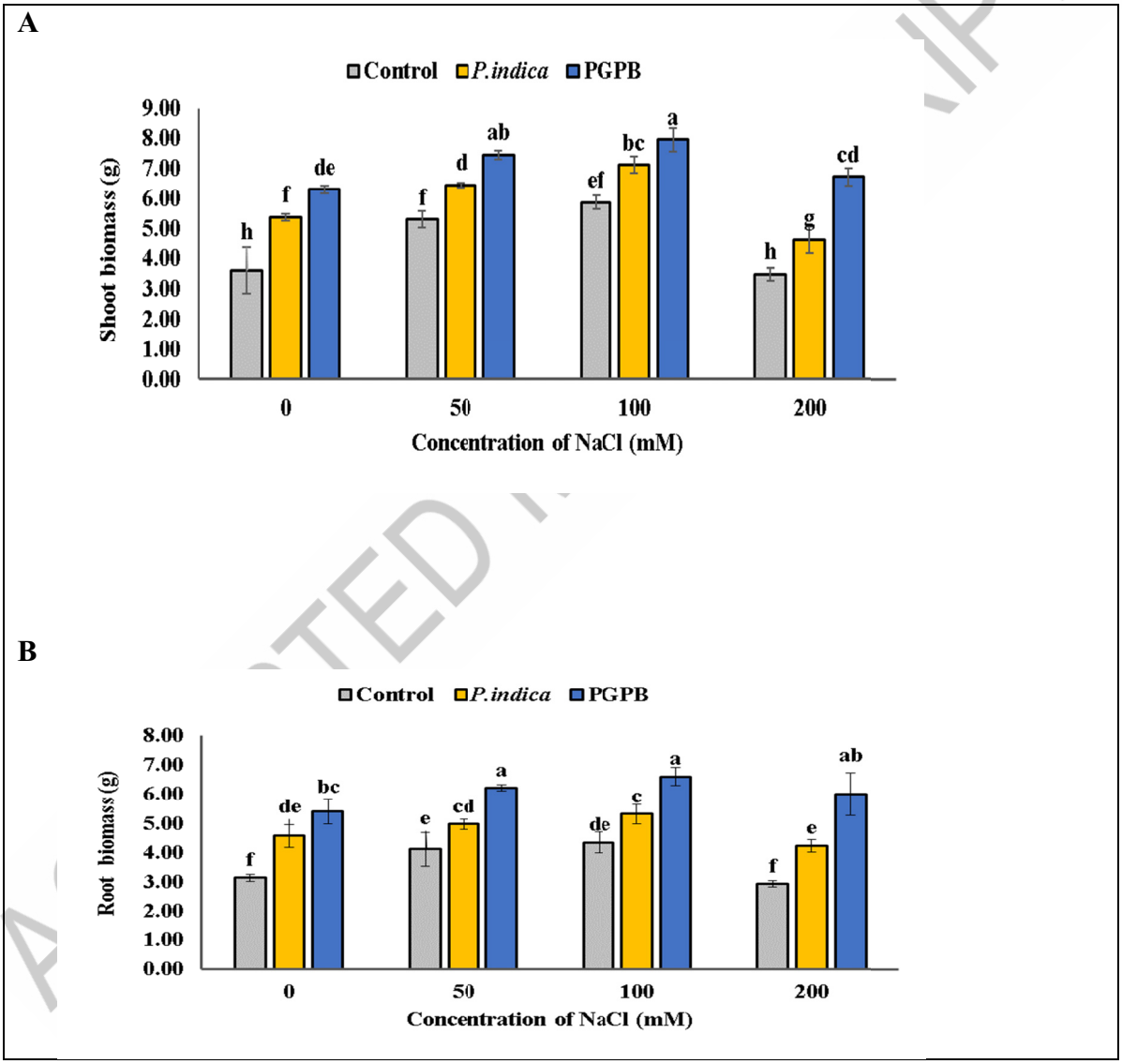


Figure 2 Effects of NaCl and microbial inoculation in HD 2967 wheat cultivar on (**A**) shoot biomass and; (**B**) root biomass. Values represent means ± S.D. (n = 3). Different letters indicate statistically significant differences for $p < 0.05$.

3.2. Oxidative Damage

Under saline conditions, MDA content in the leaves showed a rectilinear increase with increasing levels of salinity in microbial inoculated and uninoculated HD 2967 wheat cultivar. On the whole, PGPB and *P. indica* inoculated plants showed lower level of MDA than uninoculated control plants at each levels of salinity (**Figure 3 A**). Noteworthy was that there was no noticeable significant difference in lipid peroxidation level between response of PGPB and *P. indica* at 0 and 200 mM NaCl treatments.

LOX activity in the leaves of microbial inoculated and uninoculated plants increased with increased level of salinity; however, a significant difference was recorded between microbial inoculated and uninoculated plants (**Figure 3 B**). LOX activity in microbial inoculated plants was notably lower at each salinity level. However, PGPB inoculated plants showed lesser activity of LOX than that of the *P. indica*, at all levels of salinity.

As the levels of salinity increased, there was a gradual increase in the relative permeability of cell membrane in microbial inoculated and uninoculated wheat plants (**Figure 3 C**). There was no significant difference in relative permeability of membrane between *P. indica* inoculated plants and uninoculated plants at 0,50,100 mM NaCl treatments; however, at 200 mM NaCl *P. indica* inoculated plants showed lower relative permeability of membrane than uninoculated plants. At 0 mM NaCl treatment PGPB inoculated plants showed no significant difference in relative permeability of membrane as compared to uninoculated plants; however, at 50, 100 and 200 mM NaCl treatments the relative permeability was significantly lower in PGPB inoculated plants as compared to uninoculated plants. There was no significant difference in relative permeability of membrane in PGPB and *P. indica* inoculated plants at 0, 100 and 200 mM NaCl treatments; however, at 50 mM NaCl PGPB inoculated plants showed lower relative permeability of membrane than *P. indica* inoculated plants.

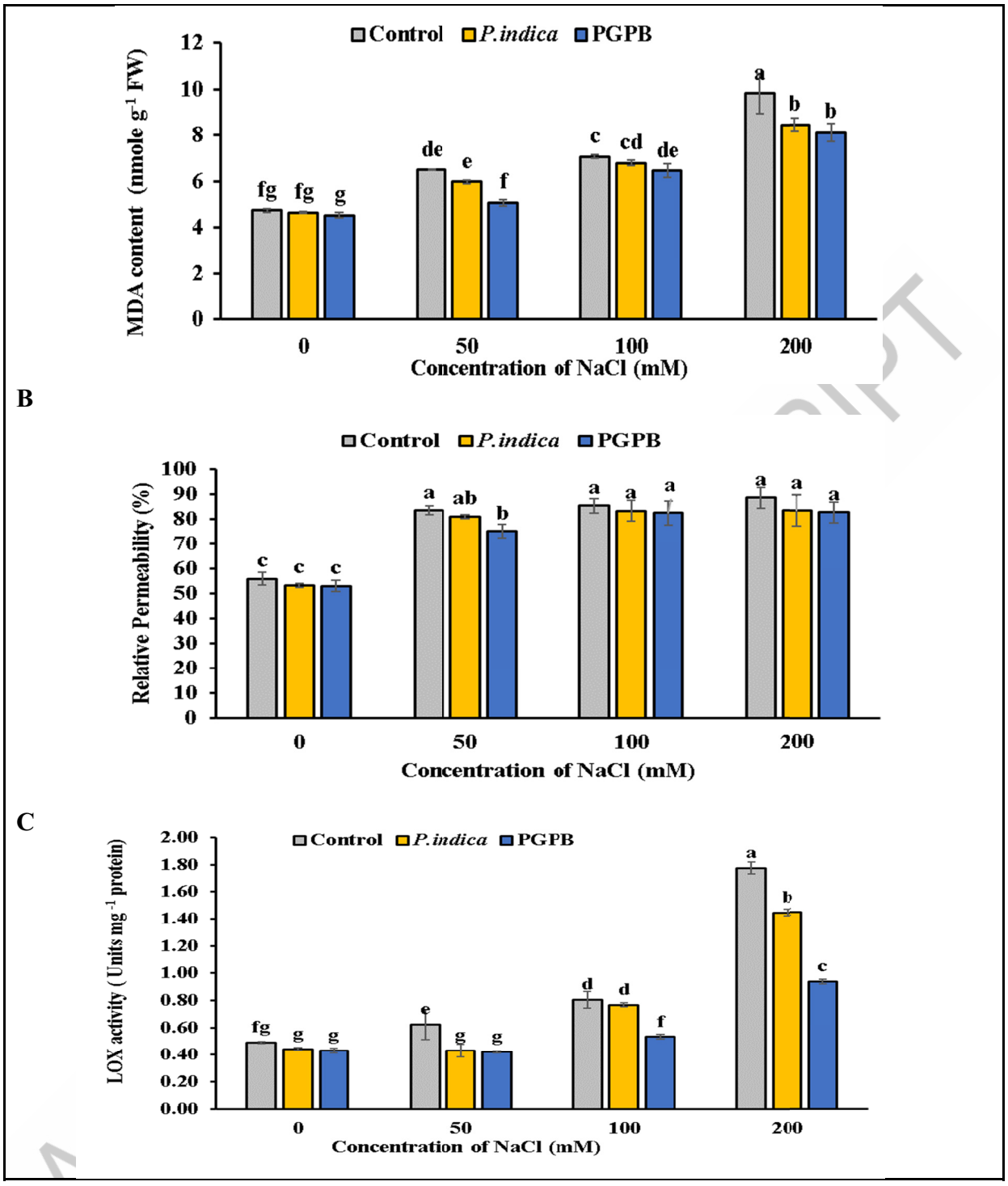
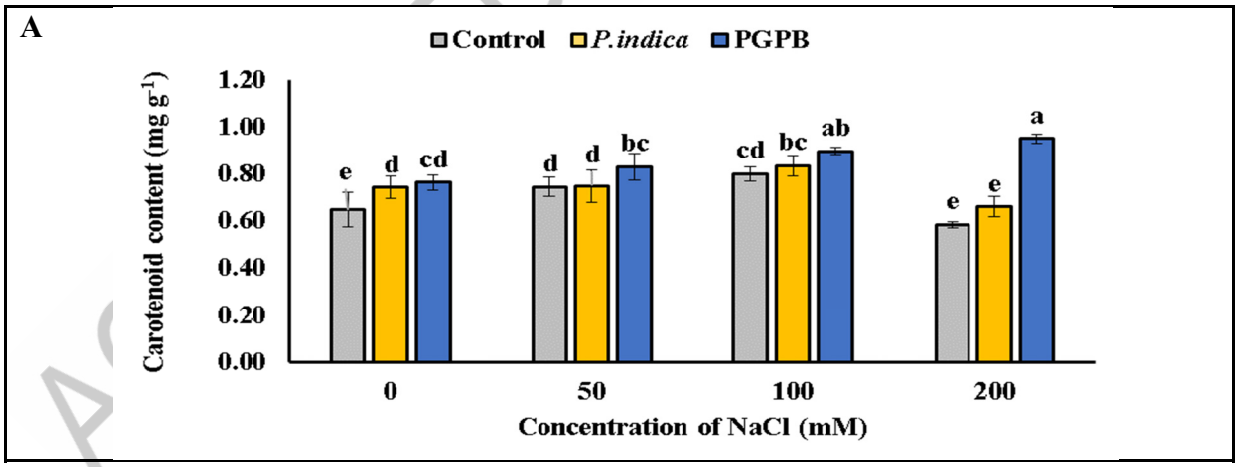


Figure 3 Effects of NaCl and microbial inoculation in HD 2967 wheat cultivar on (**A**) Malondialdehyde content; (**B**) LOX activity and; (**C**) Relative permeability. Values represent means ± S.D. (n = 3). Different letters indicate statistically significant differences for $p < 0.05$.

3.3. Antioxidant Molecules

PGPB and *P. indica* inoculation enhanced the antioxidant molecules content viz carotenoids, α-tocopherol, and proline in HD 2967 wheat plant. The carotenoid content remained unchanged at 0 and 200 mM NaCl treatments in uninoculated control plants and there was no significant difference in carotenoid content at 50 and 100 mM NaCl treatments. There was no difference in the carotenoid content between *P. indica* inoculated and uninoculated plants at 50, 100, 200 mM NaCl salinity levels. The carotenoid content was remarkably higher in PGPB inoculated plants as compared to uninoculated plants at aal levels of salinity. At 0 and 100 mM NaCl treatment there was no difference in carotenoid content between PGPB and *P. indica* inoculated plants. The content of proline and α-tocopherol were higher in microbial inoculated than uninoculated plants at all salinity levels (**Figure 4 A-C**). Accumulation of carotenoids, proline and α-tocopherol significantly increased with increasing levels of salinity (0, 50 and 100 mM NaCl); however, their content decreased when wheat plants subjected to 200 mM NaCl stress in both microbial inoculated and uninoculated plants. However, the accumulation of carotenoids and α-tocopherol was significantly greater in both PGPB and *P. indica* inoculated plants even when not subjected to salinity stress as compared to uninoculated plants.

.



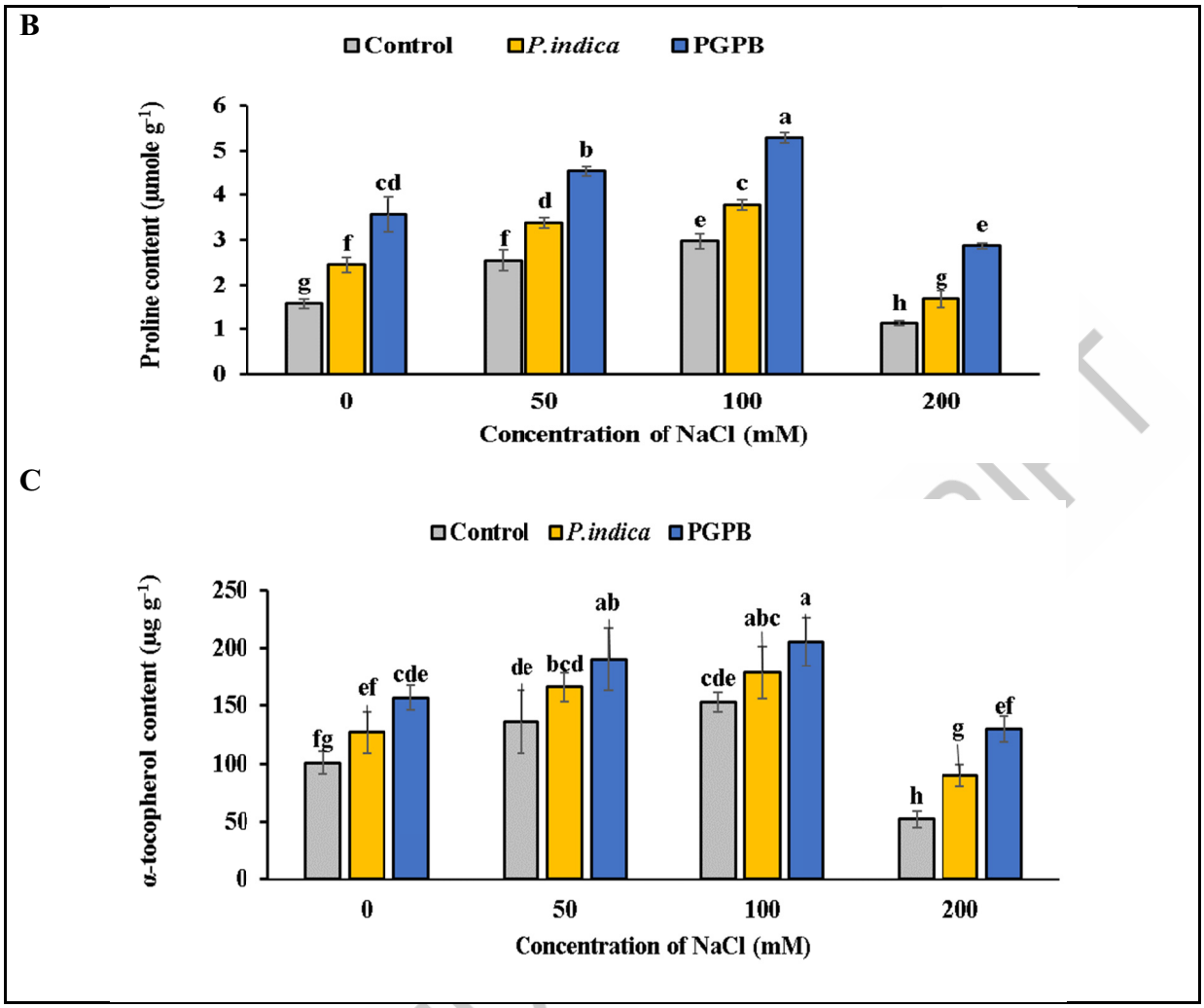


Figure 4 Effects of NaCl and microbial inoculation in HD 2967 wheat cultivar on content of (**A**) carotenoids; (**B**) proline and; (**C**) α-tocopherol. Values represent means ± S.D. (n = 3). Different letters indicate statistically significant differences for $p < 0.05$.

3.4. Antioxidant Enzyme Activity

The activities of antioxidant enzymes (SOD, CAT, and APX) were evaluated in microbial inoculated and uninoculated HD 2967 wheat plants at different levels of salinity (**Figure 5 A-C**). The activity of CAT and APX were higher in plants under salinity stress when inoculated with PGPB and *P. indica* as compared to uninoculated plants (**Figure 5 A-B**). Although microbial inoculated and uninoculated plant showed a significant increase in CAT and APX with increasing levels of NaCl (0, 50, and 100 mM NaCl); however, at 200 mM NaCl level their activity is declined. The SOD activity remained unchanged at 0 and 200 mM NaCl treatments in uninoculated control plants and there was no significant difference in SOD activity at 50 and 100 mM NaCl salinity level in uninoculated plants. In *P. indica* inoculated

plants SOD activity decreased with increase of salinity level at the activity remained same at 0 and 50 mM NaCl treatments and 100 and 200 mM NaCl treatments. In PGPB inoculated plants there was high SOD activity at 0, 100, 200 mM NaCl salinity level as compared to uninoculated plants; however, activity remained same at 50 mm NaCl salinity level. The SOD activity was higher in PGPB inoculated plants as compared to *P. indica* inoculated plants at all levels of salinity.

.

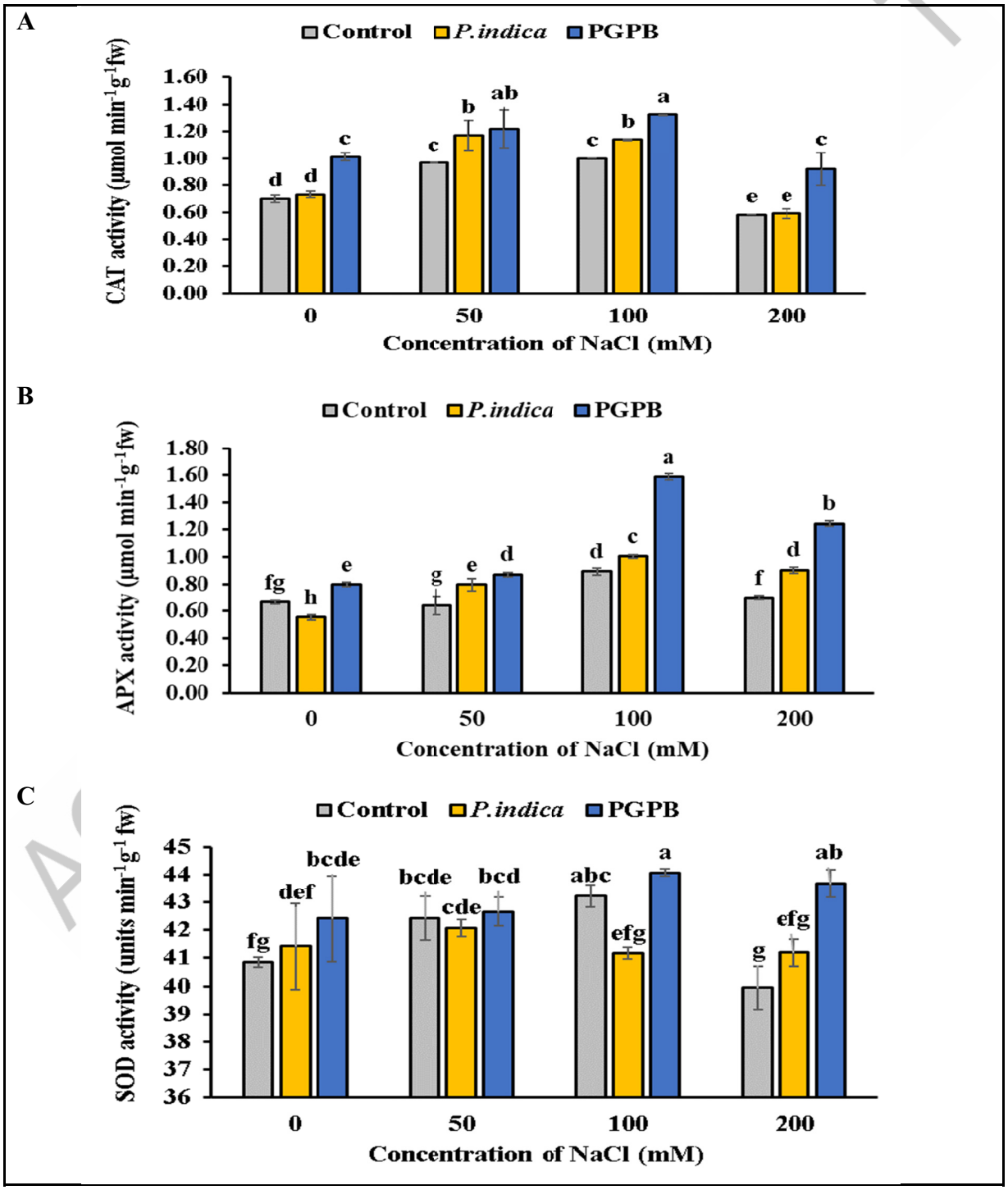


Figure 5 Effects of NaCl and microbial inoculated in HD 2967 wheat cultivar on activity of (**A**) CAT; (**B**) APX and; (**C**) SOD. Values represent means ± S.D. (n = 3). Different letters indicate statistically significant differences for $p < 0.05$.

4. Discussion

The crop plants subjected to salinity stress has deleterious effects on various biological processes via generation of ROS and thereby reduces plant growth, productivity and yield. Excessive production of ROS damages various cellular components such as nucleic acid, lipid and proteins, impacting the cellular integrity, which could lead to cell death. Salt stress has a deleterious effect on plant biomass production, enhance the production of ROS, and increase ion toxicity that limits plant cell functioning [47]. Soil microorganisms can be explored as plant growth stimulators and to minimizing deleterious effect of salinity via modulation of enzymatic and non-enzymatic antioxidant molecules accumulation. Present study demonstrated the efficacy of PGPB and *P. indica* in ameliorating salinity-induced oxidative damage in HD 2967 wheat cultivar via production of enzymatic and non-enzymatic antioxidant molecules. Findings in the present study showed a substantial increase in the plant growth as a result of PGPB and *P. indica* inoculation. Enhanced plant growth and yield by PGPB in wheat and *P. indica* in rice have also been previously reported [27,48].

Our results showed that shoot and root biomass was significantly reduced at 200 mM NaCl treatment in both microbial inoculated and uninoculated plants. The decrease in shoot and root dry weight was reported in maize varieties subjected to salt stress [49], which might be due to the production of ROS and increased ion toxicity [47]. The reduced biomass might be due to decreased accumulation non-enzymatic antioxidants (proline, carotenoids, α-tocopherol), decreased activity of enzymatic antioxidants (CAT, APX and SOD) and increased relative permeability, MDA content and LOX activity as reported in this investigation. However, higher shoot and root dry weight in PGPB and *P. indica* inoculated plants shows the positive effect of microbes on tolerance of HD 2967 under salinity stress. The effect of *P. indica* inoculation on enhanced biomass production has also been reported in

ACCEPTED MANUSCRIPT

barley [30], *Arabidopsis* [50] and maize [51] under salt stress. Microbial inoculation displayed an increase in dry weight of shoot and root at different levels of salinity which might be due to the increased accumulation of osmolytes, carotenoids, reduced ion leakage. These results are in accordance with the earlier findings of inoculation of PGPR in wheat [52] and *Zea mays* [53].

Salinity leads to production of excessive free radicals responsible for oxidative damage, and reduced plant growth [54,55]. Further, the overproduction of ROS leads to oxidation of polysaturated fatty acids of the membrane lipid layer, which is more likely under saline conditions. The MDA, which is a product of peroxidation of lipids, is considered as an indicator of oxidative damage induced by salt stress. Lipid peroxidation largely influence the ultrastructure of cellular membrane, alter their permeability, and inactivate membrane-linked receptors [10]. This study revealed that salinity-induced ROS generation leads to the higher level of MDA content in both microbial inoculated and uninoculated plants at all salinity levels. Increased LOX activity also increased lipid peroxidation level in plants as reported in this study. Plants subjected to 200 mM NaCl accrued more MDA content and membrane damage. This observation supports previous findings which showed that salinity stress induces enhanced production of MDA [56], thereby affecting the integrity of membrane [57]. However, microbial inoculation showed lower MDA content and relative permeability of membrane (**Figure 3 C**) than uninoculated plants at all levels of salinity, which indicates that PGPB and *P. indica* inoculation reduces oxidative damage and maintains the integrity of cellular membrane. The PGPB and *P. indica* were able to maintain lower MDA content by activating the adaptive responses such as synthesis of non-enzymatic molecules (carotenoids, proline and α-tocopherol) and enzymatic antioxidants (CAT, APX and SOD). These findings are similar to previous findings on improved water stress tolerance of *P. indica* inoculated rice by maintain low MDA content through enhanced ROS scavenging system [58]. This observation is in accordance with earlier report which shows that *Bacillus aquimaris* inoculation in *Zea mays* [59] and AMF inoculation in tomato [60] reduced the level of lipid peroxidation and ion leakage.

There was a remarkable difference membrane leakage between inoculated and uninoculated plants. The inoculated plants had lower leakage as compare to uninoculated plants. Thereby it shows that wheat plants on inoculation could have better management of cellular membrane

ACCEPTED MANUSCRIPT

integrity and its permeability. Furthermore, lower amount of MDA content, LOX activity and higher content of carotenoids, proline and α-tocopherol in inoculated plants confirms that microbial inoculation provide better safeguard under salt stress conditions. These results support previous observation of such lower MDA level and LOX activity in *Brevibacterium linens* inoculated *Oryza sativa* plant subjected to salinity stress [61]. Waller et al. [29] reported that there was enhanced production of antioxidants in *P. indica* inoculated barley, which reduces the salinity-induced lipid peroxidation. However, these findings do not corroborate the previous findings of Kumari et al. [62] where there was no correlation between lipid peroxidation level and LOX activity in bacterial colonized soyabean plants under salt stress.

LOX enzyme converts lipids into hydroperoxyl fatty acids [12]. Salinity induced enhanced LOX activity is linked with the production of free radicals and peroxidation of lipids, causing damage to the cell. LOX activity showed a rectilinear increase with the increase in salinity level in microbial inoculated and uninoculated plants. However, compared to uninoculated, LOX activity in microbial inoculated plants was notably lower at each level of NaCl treatment thus, reduces lipid peroxidation, maintain membrane integrity, reduces ion leakage and enhanced plant growth. The PGPB and *P. indica* were able to maintain lower LOX activity by activating the synthesis of non-enzymatic molecules (carotenoids, proline and α-tocopherol) and enzymatic antioxidants (CAT, APX and SOD). The reduced LOX activity in microbial inoculated plants is in agreement with earlier report in rice plant [61].

The non-enzymatic antioxidants include carotenoids and proline [63]. In the present study, salinty stress led to enhanced accumulation of antioxidant molecules viz. carotenoids, proline and α-tocopherol in inoculated as compared to uninoculated plants thereby showing a better management of MDA content and cellular membrane integrity. These results support previous observation of such high accumulation of antioxidant molecules in fenugreek plants grow in a saline soil [33]. Lower level of lipid peroxidation in microbial inoculated plants could be a result of high concent of α-tocopherol synthesis, which disrupts lipid oxidation pathway [64].

Carotenoids are the harvester of light during photosynthesis. Moreover, they act as chloroplast membrane stabilizers hence reducing ion leakage and peroxidation of lipids

[21,22]. In the present study, the increased accumulation of carotenoids in PGPB and *P. indica* inoculated plants, lowered the oxidative damage, as carotenoids prevent generation of singlet oxygen [65]. Proline has a major role in osmotic adjustment, stabilizing subcellular structures, and scavenge ROS under drought stress [13]. Indeed, proline accumulation under water and salinity stress may facilitate plants to withstand such environmental stresses. The increased content of proline in PGPB and *P. indica* colonized wheat plants indicated greater salt tolerance as compared to uninoculated plants which might be due to reduced lipid peroxidation and increased membrane stability, signifying a role of proline as an indicator for salinity tolerance [66,25]. The increased proline content has also been found to alleviate oxidative stress caused by salt stress in microbial inoculated maize plants [67], *P. indica* inoculated rice plants [28] and in PGPR inoculated soybean [68].

The α-tocopherol being an antioxidant deactivates ROS generation during photosynthesis (mainly $^1O_2$ and $OH^{\bullet}$), and lowers lipid peroxidation level by scavenging lipid peroxyl radicals [34]. High content of α-tocopherol in PGPB and *P. indica* colonized plants as compared to microbial uninoculated plants confirms the role of α-tocopherol as free radical scavenger and lipid peroxidation modulator in plants during salinity stress [26,64]. These results substantiate the fact that higher content of α-tocopherol in AMF colonized fenugreek plant mitigates injurious effect of salt stress as has been reported by Evelin and Kapoor [33]. Present study revealed that inoculation of wheat plants by PGPB and *P. indica* enhanced salt tolerance in wheat via regulating cellular enzymatic and non-enzymatic antioxidant system.

There was a remarkable difference between antioxidant enzymes activity of microbial inoculated and uninoculated plants. At each level of salinity, microbial inoculation exhibited better plant growth than uninoculated plants, that may be due to increased activity of ROS scavenging antioxidant enzymes (CAT, APX, and SOD). Both PGPB and *P. indica* inoculated plant exhibited higher activities of SOD as compared to uninoculated plants under salinity stress. SOD has a significant role in protecting plant from oxidative damage by converting $O_2^{\cdot-}$ to H_2O_2. H_2O_2 generated needs to be detoxified by conversion to H_2O. Further, the speedy scavenging of H_2O_2 is done by enhanced activities of CAT, and APX, which catalyze the breakdown of H_2O_2 produced by SOD thus protecting the cell from oxidative damage. At all salinity level, the CAT, and APX activities were more in microbial inoculated plants than uninoculated plants. CAT and APX activity increased with increasing salinity levels (0, 50, and 100mM NaCl), however, their activity declined at 200 mM NaCl treatment. This may be attributed due to following factors: (i) salinity stress aggravates CAT

ACCEPTED MANUSCRIPT

destruction by proteinases, (ii) decrease iron content under salinity stress [69,70], and (iii) synthesis of new enzymes are prevented by salinity stress [71]. Overall, the present study exhibited lower oxidative stress in PGPB and *P. indica* that may be a result of i) improved membrane stability by lowered lipid peroxidation, loss of ions and LOX enzyme activity, and ii) a better antioxidative capacity of PGPB and *P. indica* inoculated than uninoculated plants.

In conclusion, PGPB and *P. indica* inoculation enhanced production of non-enzymatic antioxidant molecules and activities of antioxidant enzymes at all levels of salinity. Therefore, PGPB and *P. indica* inoculated HD 2967 wheat cultivar possessed better antioxidant defence system to prevent oxidative damage, and therefore more tolerant to salinity stress. The results presented in this study clearly indicated that PGPB and *P. indica* improved salt stress tolerance potential of HD 2967 wheat plants, by down regulating the ROS response due to improved enzymatic and non-enzymatic antioxidant system. The enhanced level of antioxidant molecules in PGPB and *P. indica* inoculated plants under salinity stress, also substantiate the supposition that plant-microbe interaction could ameliorate salinity stress and tolerance to salt stress in HD 2967 wheat cultivar.

Acknowledgments: MS is thankful to Central Laboratory for Soil & Plant Analysis, Division of SS & AC, ICAR-IARI, New Delhi for soil testing and Department of Microbiology, IARI, New Delhi for providing PGPBs inoculum. Author is also thankful to *Professor* Ajit Verma, Department of Microbiology, Amity University, Noida, Uttar Pradesh, India for providing *Piriformospora indica* inoculum.

Conflicts of Interest: Authors declares no conflict of interest.

Funding: There is no funding to this research

References

1. Zhang, S.; Fan, C.; Wang, Y.; Xia, Y.; Xiao, W.; Cui, X. Salt-Tolerant and Plant Growth-Promoting Bacteria Isolated from High-Yield Paddy Soil. *Can. J. Microbiol.* **2018,** *64* (12), 968–978. doi:10.1139/cjm-2017-0571.

2. Sarkar, A.; Ghosh, P. K.; Pramanik, K.; Mitra, S.; Soren, T.; Pandey, S.; Mondal, M. H.; Maiti, T. K. A Halotolerant Enterobacter sp. Displaying ACC Deaminase Activity Promotes Rice Seedling Growth Under Salt Stress. *Res. Microbiol.* **2018,** *169* (1), 20–32. doi:10.1016/j.resmic.2017.08.005.

3. Shrivastava, P.; Kumar, R. Soil Salinity: A Serious Environmental Issue and Plant Growth Promoting Bacteria as One of the Tools for Its Alleviation. *Saudi J. Biol. Sci.* **2015,** *22* (2), 123–131. doi:10.1016/j.sjbs.2014.12.001.

4. Jamil, A.; Riaz, S.; Ashraf, M.; Foolad, M. R. Gene Expression Profiling of Plants Under Salt Stress. *Crit. Rev. Plant Sci.* **2011,** *30* (5), 435–458. doi:10.1080/07352689.2011.605739.

5. Gengmao, Z.; Quanmei, S.; Yu, H.; Shihui, L.; Changhai, W. The Physiological and Biochemical Responses of a Medicinal Plant (Salvia miltiorrhiza L.) to Stress Caused by Various Concentrations of NaCl. *PLOS ONE* **2014,** *9* (2), e89624. doi:10.1371/journal.pone.0089624.

6. Das, K.; Roychoudhury, A. Reactive Oxygen Species (ROS) and Response of Antioxidants as ROS-Scavengers During Environmental Stress in Plants. *Front. Environ. Sci.* **2014,** *2,* 53. doi:10.3389/fenvs.2014.00053.

7. Evelin, H.; Kapoor, R.; Giri, B. Arbuscular Mycorrhizal Fungi in Alleviation of Salt Stress: A Review. *Ann. Bot.* **2009,** *104* (7), 1263–1280. doi:10.1093/aob/mcp251.

8. Sharma, P.; Jha, A. B.; Dubey, R. S.; Pessarakli, M. Reactive Oxygen Species, Oxidative Damage, and Antioxidative Defense Mechanism in Plants Under Stressful Conditions. *J. Bot.* **2012**, 1-26. doi:10.1155/2012/217037.

9. Yu, D.; Boughton, B. A.; Hill, C. B.; Feussner, I.; Roessner, U.; Rupasinghe, T. W. T. Insights into Oxidized Lipid Modification in Barley Roots as an Adaptation Mechanism to Salinity Stress. *Front. Plant Sci.* **2020,** *11,* 1. doi:10.3389/fpls.2020.00001.

10. Mansour, M. M. F.; Salama, K. H. A. Cellular Basis of Salinity Tolerance in Plants. *Environ. Exp. Bot.* **2004,** *52* (2), 113–122. doi:10.1016/j.envexpbot.2004.01.009.

ACCEPTED MANUSCRIPT

11. Lim, C. W.; Han, S. W.; Hwang, I. S.; Kim, D. S.; Hwang, B. K.; Lee, S. C. The Pepper Lipoxygenase CaLOX1 Plays a Role in Osmotic, Drought and High Salinity Stress Response. *Plant Cell Physiol.* **2015,** *56* (5), 930–942. doi:10.1093/pcp/pcv020.

12. Blée, E. Impact of Phyto-Oxylipins in Plant Defense. *Trends Plant Sci.* **2002,** *7* (7), 315–322. doi:10.1016/s1360-1385(02)02290-2.

13. Ashraf, M.; Foolad, M. R. Roles of Glycine Betaine and Proline in Improving Plant Abiotic Stress Resistance. *Environ. Exp. Bot.* **2007,** *59* (2), 206–216. doi:10.1016/j.envexpbot.2005.12.006.

14. Saxena, S. C.; Joshi, P. K.; Grimm, B.; Arora, S. Alleviation of Ultraviolet-C Induced Oxidative Through Overexpression of Cytosolic Ascorbate Peroxidase. *Biologia* **2011,** *66* (6), 1052–1059. doi:10.2478/s11756-011-0120-4.

15. Sadiq, M.; Akram, N. A.; Ashraf, M.; Al-Qurainy, F.; Ahmad, P. Alpha-Tocopherol-Induced Regulation of Growth and Metabolism in Plants Under Non-Stress and Stress Conditions. *J. Plant Growth Regul.* **2019,** *38* (4), 1325–1340. doi:10.1007/s00344-019-09936-7.

16. Qian, H. F.; Peng, X. F.; Han, X.; Ren, J.; Zhan, K. Y.; Zhu, M. The Stress Factor, Exogenous Ascorbic Acid, Affects Plant Growth and The Antioxidant System in *Arabidopsis Thaliana*. *Russ. J. Plant Physiol.* **2014**, 61, 467–475. doi: 10.1134/S1021443714040141.

17. Naz, H.; Akram, N. A.; Ashraf, M. Impact of Ascorbic Acid on Growth and Some Physiological Attributes of Cucumber (*Cucumis Sativus*) Plants Under Water-Deficit Conditions. *Pak. J. Bot*. **2016**, 48(3), 877–883.

18. Agami, R. A. Applications of Ascorbic Acid or Proline Increase Resistance to Salt Stress in Barley Seedlings. *Biol. Plant.* **201**4, 58, 341–347. doi: 10.1007/s10535-014-0392-y.

ACCEPTED MANUSCRIPT

19. Şirin, S.; Aslım, B. Determination of Antioxidant Capacity, Phenolic Acid Composition and Antiproliferative Effect Associated With Phenylalanine Ammonia Lyase (PAL) Activity in Some Plants Naturally Growing Under Salt Stress. *Medicinal Chemistry Research*. **2019**, *28*(3), 229–238.doi:10.1007/s00044-018-2278-6.

20. Tuladhar, P.; Sasidharan, S.; Saudagar, P. Role of Phenols and Polyphenols in Plant Defense Response to Biotic and Abiotic Stresses. In *Biocontrol Agents and Secondary Metabolites*, Elsevier , **2021**, 419–441. https://doi.org/10.1016/b978-0-12-822919-4.00017-x.

21. Havaux, M. Carotenoids as Membrane Stabilizers in Chloroplasts. *Trends Plant Sci.* **1998,** *3* (4)*,* 147–151. doi:10.1016/S1360-1385(98)01200-X.

22. Johnson, M. P.; Havaux, M.; Triantaphylidès, C.; Ksas, B.; Pascal, A. A.; Robert, B.; Davison, P. A.; Ruban, A. V.; Horton, P. Elevated Zeaxanthin Bound to Oligomeric LHCII Enhances the Resistance of Arabidopsis to Photooxidative Stress by a Lipid-Protective, Antioxidant Mechanism. *J. Biol. Chem.* **2007,** *282* (31)*,* 22605–22618. doi:10.1074/jbc.M702831200.

23. Alscher, R. G.; Erturk, N.; Heath, L. S. Role of Superoxide Dismutases (SODs) in Controlling Oxidative Stress in Plants. *J. Exp. Bot.* **2002,** *53* (372)*,* 1331–1341. doi:10.1093/jxb/53.372.1331.

24. Jaleel, C. A.; Riadh, K.; Gopi, R.; Manivannan, P.; Inès, J.; Al-Juburi, H. J.; Chang-Xing, Z.; Hong-Bo, S.; Panneerselvam, R. Antioxidant Defense Responses: Physiological Plasticity in Higher Plants Under Abiotic Constraints. *Acta Physiol. Plant.* **2009,** *31* (3)*,* 427–436. doi:10.1007/s11738-009-0275-6.

25. Khan, A.; Sirajuddin, Z.; Zhao, X. Q.; Javed, M. T.; Khan, K. S.; Bano, A.; Shen, R. F.; Masood, S. *Bacillus pumilus* Enhances Tolerance in Rice (*Oryza sativa* L.) to Combined Stresses of NaCl and High Boron Due to Limited Uptake of Na+. *Environ. Exp. Bot.* **2016,** *124,* 120–129. doi:10.1016/j.envexpbot.2015.12.011.

ACCEPTED MANUSCRIPT

26. Zhang, S.; Gan, Y.; Xu, B. Application of Plant-Growth Promoting Fungi Trichoderma longibrachiatum T6 Enhances Tolerance of Wheat to Salt Stress Through Improvement of Antioxidative Defense System and Gene Expression. *Front. Plant Sci.* **2016,** *7,* 1405. doi:10.3389/fpls.2016.01405.

27. Jogawat, A.; Saha, S.; Bakshi, M.; Dayaman, V.; Kumar, M.; Dua, M.; Varma, A.; Oelmüller, R.; Tuteja, N.; Johri, A. K. Piriformospora indica Rescues Growth Diminution of Rice Seedlings During High Salt Stress. *Plant Signal. Behav.* **2013,** *8* (10), e26891. doi: 10.4161/psb.26891.

28. Husaini, A. M.; Abdin, M. Z.; Khan, S.; Xu, Y. W.; Aquil, S.; Anis, M. Modifying Strawberry for Better Adaptability to Adverse Impact of Climate Change. *Curr. Sci.* **2012,** *102,* 1660–1673.

29. Waller, F.; Achatz, B.; Baltruschat, H.; Fodor, J.; Becker, K.; Fischer, M.; Heier, T.; Hückelhoven, R.; Neumann, C.; von Wettstein, D.; Franken, P.; Kogel, K. H. The Endophytic Fungus Piriformospora indica Reprograms Barley to Salt-Stress Tolerance, Disease Resistance, and Higher Yield. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102* (38)*,* 13386–13391. https://doi.org/10.1073/pnas.0504423102.

30. Islam, F.; Yasmeen, T.; Arif, M. S.; Ali, S.; Ali, B.; Hameed, S.; Zhou, W. Plant Growth Promoting Bacteria Confer Salt Tolerance in Vigna radiata by Up-Regulating Antioxidant Defense and Biological Soil Fertility. *Plant Growth Regul.* **2016,** *80* (1)*,* 23–36. doi:10.1007/s10725-015-0142-y.

31. Rojas-Tapias, D.; Moreno-Galván, A.; Pardo-Díaz, S.; Obando, M.; Rivera, D.; Bonilla, R. Effect of Inoculation with Plant Growth-Promoting Bacteria (PGPB) on Amelioration of Saline Stress in Maize (Zea mays). *Appl. Soil Ecol.* **2012,** *61,* 264–272. doi:10.1016/j.apsoil.2012.01.006.

32. Kasotia, A.; Varma, A.; Choudhary, D. K. Pseudomonas-Mediated Mitigation of Salt Stress and Growth Promotion in Glycine max. *Agric. Res.* **2015,** *4* (1)*,* 31–41. doi:10.1007/s40003-014-0139-1.

ACCEPTED MANUSCRIPT

33. Evelin, H.; Kapoor, R. Arbuscular Mycorrhizal Symbiosis Modulates Antioxidant Response in Salt-Stressed Trigonella foenum-Graecum Plants. *Mycorrhiza* **2014,** *24* (3)*,* 197–208. doi:10.1007/s00572-013-0529-4.

34. Munné-Bosch, S. The Role of Alpha-Tocopherol in Plant Stress Tolerance. *J. Plant Physiol.* **2005,** *162* (7)*,* 743–748. doi:10.1016/j.jplph.2005.04.022.

35. Ha-Tran, D. M.; Nguyen, T. T. M.; Hung, S. H.; Huang, E.; Huang, C. C. Roles of Plant Growth-Promoting Rhizobacteria (PGPR) in Stimulating Salinity Stress Defense in Plants: A Review. *Int. J. Mol. Sci.* **2021,** *22* (6), 3154. doi:10.3390/ijms22063154.

36. Xu, L.; Wang, A.; Wang, J.; Wei, Q.; Zhang, W. *Piriformospora indica* Confers Drought Tolerance on Zea mays L. Through Enhancement of Antioxidant Activity and Expression of Drought-Related Genes. *Crop J.* **2017,** *5* (3)*,* 251–258. doi:10.1016/j.cj.2016.10.002.

37. Zwiazek, J. J.; Blake, T. J. Early Detection of Membrane Injury in Black Spruce (*Picea mariana*). *Can. J. Forest Res.* **1991,** *21* (3)*,* 401–404. doi:10.1139/x91-050.

38. Heath, R. L.; Packer, L. Photoperoxidation in Isolated Chloroplasts: I. Kinetics and Stoichiometry of Fatty Acid Peroxidation. *Arch. Biochem. Biophys.* **1968,** *125* (1)*,* 189–198. doi:10.1016/0003-9861(68)90654-1.

39. Doderer, A.; Kokkelink, I.; van der Veen, S.; Valk, B. E.; Schram, A. W.; Douma, A. C. Purification and Characterization of Two Lipoxygenase Isoenzymes from Germinating Barley. *Biochim. Biophys. Acta* **1992,** *1120* (1)*,* 97–104. doi:10.1016/0167-4838(92)90429-H.

40. Bates, L. S.; Waldren, R. P.; Teare I. D. Rapid Determination of Free Proline for Water Stress Studies. *Plant Soil* **1973,** *39,* 205–207.

41. Sadasivam, S.; Manickam, A. Biochemical Methods. *Age*; New: New Delhi **2008**.

42. Hiscox, J. D.; Israelstam, G. F. A Method for the Extraction of Chlorophyll from Leaf Tissue Without Maceration. *Can. J. Bot.* **1979,** *57* (12)*,* 1332–1334. doi:10.1139/b79-163.

ACCEPTED MANUSCRIPT

43. Arnon, D. I. Copper Enzymes Isolated Chloroplasts, Polyphenoloxidase in Beta vulgaris. *Plant Physiol.* **1949,** *24* (1)*,* 1–15. doi:10.1104/pp.24.1.1.

44. Dhindsa, R. S.; Plumb-Dhindsa, P.; Thorpe, T. A. Leaf Senescence: Correlated with Increased Levels of Membrane Permeability and Lipid Peroxidation, and Decreased Levels of Superoxide Dismutase and Catalase. *J. Exp. Bot.* **1981,** *32* (1)*,* 93–101. doi:10.1093/jxb/32.1.93.

45. Teranishi, Y.; Tanaka, A.; Osumi, M.; Fukui, S. Catalase Activities of Hydrocarbon-Utilizing *Candida* Yeasts. *Agric. Biol. Chem.* **1974,** *38* (6)*,* 1213–1220. doi:10.1080/00021369.1974.10861301.

46. Nakano, Y.; Asada, K. Hydrogen Peroxide Is Scavenged by Ascorbate-Specific Peroxidase in Spinach-Chloroplasts. *Plant Cell Physiol.* **1981,** *22,* 867–880. https://doi.org/10.1093/oxfordjournals.pcp.a076232.

47. Khan, N. A.; Khan, M. I. R.; Asgher, M.; Fatma, M.; Masood, A.; Syeed, S. Salinity Tolerance in Plants: Revisiting the Role of Sulphur Metabolites. *J. Plant Biochem. Physiol.* **2014,** *2,* 120. DOI: 10.4172/2329-9029.1000120.

48. Barra, P. J.; Inostroza, N. G.; Acuña, J. J.; Mora, M. L.; Crowley, D. E.; Jorquera, M. A. Formulation of Bacterial Consortia from Avocado (Persea americana Mill.) and Their Effect on Growth, Biomass and Superoxide Dismutase Activity of Wheat Seedlings Under Salt Stress. *Appl. Soil Ecol.* **2016,** *102,* 80–91. doi:10.1016/j.apsoil.2016.02.014.

49. Eker, S.; Cömertpay, G.; Konuskan, O.; Ulger, A. C.; Ozt¨urk, L.; *Cakmak*, I. Effect of salinity stress on dry matter production and ion accumulation in hybrid maize varieties. *Turk J Agric For.* **2006,** *30* (5)*,* 365–373.

50. Abdelaziz, M. E.; Kim, D.; Ali, S.; Fedoroff, N. V.; Al-Babili, S. The Endophytic Fungus *Piriformospora indica* Enhances *Arabidopsis thaliana* Growth and Modulates Na^+/K^+ Homeostasis Under Salt Stress Conditions. *Plant Sci.* **2017,** *263,* 107–115. doi:10.1016/j.plantsci.2017.07.006.

ACCEPTED MANUSCRIPT

51. Yun, P.; Xu, L.; Wang, S.; Shabala, L.; Shabala, S.; Zhang, W. Piriformospora indica Improves Salinity Stress Tolerance in Zea mays L. Plants by Regulating Na+ and K+ Loading in Root and Allocating K+ in Shoot. *Plant Growth Regul.* **2018,** *86* (2), 323–331. doi:10.1007/s10725-018-0431-3.

52. Upadhyay, S. K.; Singh, J. S.; Singh, D. P. Exopolysaccharide-Producing Plant Growth-Promoting Rhizobacteria Under Salinity Condition. *Pedosphere* **2011,** *21* (2), 214–222. doi:10.1016/S1002-0160(11)60120-3.

53. Misra, S.; Chauhan, P. S. ACC Deaminase-Producing Rhizosphere Competent Bacillus spp. Mitigate Salt Stress and Promote Zea mays Growth by Modulating Ethylene Metabolism. *3 Biotech* **2020,** *10* (3), 119. doi:10.1007/s13205-020-2104-y.

54. Mittler, R. Oxidative Stress, Antioxidants and Stress Tolerance. *Trends Plant Sci.* **2002,** *7* (9), 405–410. doi:10.1016/s1360-1385(02)02312-9.

55. Chang, W.; Sui, X.; Fan, X. X.; Jia, T. T.; Song, F. Q. Arbuscular Mycorrhizal Symbiosis Modulates Antioxidant Response and Ion Distribution in Salt-Stressed Elaeagnus angustifolia Seedlings. *Front. Microbiol.* **2018,** *9,* 652. doi:10.3389/fmicb.2018.00652.

56. Wi, S. G.; Chung, B. Y.; Kim, J. -H.; Lee, K. -S.; Kim, J. -S. Deposition Pattern of Hydrogen Peroxide in the Leaf Sheaths of Rice Under Salt Stress. *Biol. Plant.* **2006,** *50* (3), 469–472. doi:10.1007/s10535-006-0073-6.

57. Maccarrone, M.; Melino, G.; Finazzi-Agrò, A. Lipoxygenases and Their Involvement in Programmed Cell Death. *Cell Death Differ.* **2001,** *8* (8), 776–784. doi:10.1038/sj.cdd.4400908.

58. Tsai, H. J.; Shao, K. H.; Chan, M. T.; Cheng, C. P.; Yeh, K. W.; Oelmüller, R.; Wang, S. J. *Piriformospora indica* Symbiosis Improves Water Stress Tolerance of Rice Through Regulating Stomata Behavior and ROS Scavenging Systems. *Plant Signal. Behav.* **2020,** *15* (2), 1722447. doi:10.1080/15592324.2020.1722447.

ACCEPTED MANUSCRIPT

59. Li, H. Q.; Jiang, X. W. Inoculation with Plant Growth-Promoting Bacteria (PGPB) Improves Salt Tolerance of Maize Seedling. *Russ. J. Plant Physiol.* **2017,** *64* (2)*,* 235–241. doi:10.1134/S1021443717020078.

60. Hajiboland, R.; Aliasgharzadeh, N.; Laiegh, S. F.; Poschenrieder, C. Colonization with Arbuscular Mycorrhizal Fungi Improves Salinity Tolerance of Tomato (*Solanum lycopersicum* L.) Plants. *Plant Soil* **2010,** *331* (1–2)*,* 313–327. doi:10.1007/s11104-009-0255-z.

61. Chatterjee, P.; Kanagendran, A.; Samaddar, S.; Pazouki, L.; Niinemets, Ü. Brevibacterium linens RS16 Inoculation Reduces Salt Stress-Induced Foliar Volatileemission in Oryza Sativagenotypes. *Microbiol. Res.* **2018,** *215,* 89–101–101. doi:10.1016/j.micres.2018.06.007.

62. Kumari, Sarita and Vaishnav, Anukool and JAIN, Shekhar and A, Varma and DK, Choudhary. Bacterial-mediated induction of systemic tolerance to salinity with expression of defense related proteins in soybean (Glycine max L. Merrill). *J. Plant Growth Regul.* **2015**, 34, 558-573. doi: 10.1007/s00344-015-9490-0.

63. Ahmad, P.; Jaleel, C. A.; Salem, M. A.; Nabi, G.; Sharma, S. Roles of Enzymatic and Nonenzymatic Antioxidants in Plants During Abiotic Stress. *Crit. Rev. Biotechnol.* **2010,** *30* (3)*,* 161–175. doi:10.3109/07388550903524243.

64. Serbinova, E. A.; Packer, L. Antioxidant Properties of Tocopherol and Tocotrienol. *Methods Enzymol.* **1994,** *234,* 354–366. doi:10.1016/0076-6879(94)34105-2.

65. Fyfe, P.; Cogdell, R. J.; Hunter, C. N.; Jones, M. R. Study of the Carotenoid Binding Pocket of the Photosynthetic Reaction Center from the Purple Bacterium Rhodobacter sphaeroides. In Mathis, P., Ed. *Photosynthesis from Light to Biosphere. Proceeding of 10th International Photosynthesis Congress*, Montpelleir, France, **1995***,* 47–50.

66. Younesi, O.; Moradi, A. Effects of Plant Growth-Promoting Rhizobacterium (PGPR) and Arbuscular Mycorrhizal Fungus (AMF) on Antioxidant Enzyme Activities in Salt stressed Bean (*Phaseolus vulgaris* L.). *Agriculture* **2014,** *60,* 10–21. DOI: 10.2478/agri-2014-0002.

ACCEPTED MANUSCRIPT

67. Rafiq, K.; Akram, M. S.; Shahid, M.; Qaisar, U.; Rashid, N. Enhancement of Salt Tolerance in Maize (*Zea mays* L.) Using Locally Isolated *Bacillus* sp. *Biologia* **2020,** *75* (9), 1425–1436. doi:10.2478/s11756-020-00435-9.

68. Han, H. S.; Lee, K. D. Physiological Responses of Soybean Inoculation of Bradyrhizobium japonicum PGPR in Saline Soil Conditions. *Res. J. Agric. Biol. Sci.* **2005,** *1* (3), 216–221.

69. Shim, I. S.; Momose, Y.; Yamamoto, A.; Kim, D. W.; Usui, K. Inhibition of Catalase Activity by Oxidative Stress and Its Relationship to Salicylic Acid Accumulation in Plants. *Plant Growth Regul.* **2003,** *39* (3), 285–292. doi:10.1023/A:1022861312375.

70. Foyer, C. H.; Noctor, G. Oxidant and Antioxidant Signaling in Plants: Are-Evaluation of the Concept of Oxidative Stress in a Physiological Context. *Plant Cell Environ.* **2005,** *28* (8), 1056–1071. doi:10.1111/j.1365-3040.2005.01327.x.

71. Feierabend, J.; Engel, S. Photoinactivation of Catalase In Vitro and in Leaves. *Arch. Biochem. Biophys.* **1986,** *251* (2), 567–576. doi:10.1016/0003-9861(86)90365-6.

ACCEPTED MANUSCRIPT

Proceedings of the Fifth International Conference on Trends in Electronics and Informatics (ICOEI).
IEEE Xplore Part Number:CFP21J32-ART; ISBN:978-1-6654-1571-2

Movie genres and beyond: An analytical survey of classification techniques

Avleen Singh Mehal
Dept. of Software Engineering
Delhi Technological University
New Delhi, India
avleenmehal@gmail.com

Kalpana Meena
Dept. of Software Engineering
Delhi Technological University
New Delhi, India
kalpanameena_2k17se52@dtu.ac.in

Rijul Bir Singh
Dept. of Software Engineering
Delhi Technological University
New Delhi, India
rijulbirsingh_2k17se93@dtu.ac.in

Prashant Giridhar Shambharkar
Dept. of Computer Engineering
Delhi Technological University
New Delhi, India
prashant.shambharkar@dtu.ac.in

2021 5th International Conference on Trends in Electronics and Informatics (ICOEI) | 978-1-6654-1571-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICOEI51242.2021.9453021

***Abstract*— In today's world with a plethora of options available on movie content, it is very important to have models which could classify these large contents based on their genres. Movies are the most widely entertaining source all over the world. Every person has his own taste of movies, some feel action movies are better, some would say romantic. So the online platforms must have the collection of movies according to a particular category i.e. based on genres. This paper has proposed an extensive and exhaustive survey for various models and techniques, which could be applied to solve the challenges involved in performing the genre based classification. This paper summarizes and compares all those techniques by classifying them into four broad domains -through text mining, poster detection, audio and video features of the movies. These domains included various classifying techniques: Bi-directional LSTM, gram layer, scene categorization and CNN respectively. With such an extensive amount of work going in the field in the last not many years, this review paper will help specialists in the field of movie genre classification.**

I. INTRODUCTION

Video based content assume a major function in satisfying individuals' delight needs. On account of Internet innovation headways, content consumers approach a remarkable measure of motion pictures from different online administrations.//With increasing number of users using the internet in India and around the world have shot up multiple folds in recent past, thanks to free internet and lower rate/cheap internet packs. Because of this there has been a sudden increase in number of people engaging in consumption of online entertainment channels having web shows, web series, movies, daily soaps etc. This has made the need for programmed content-driven moving-picture show suggestions of genres before a user starts watching it.

Number of interesting models has been already proposed by researchers in the recent past. In [1] algorithm of SVM has een extensively used along with text containing information about the video which helps in the classification process. Main focus in this model is given to the three basic steps for text based classification i.e. word segmentation, feature engineering, text representation, and designing followed by implementation of a movie classification system based on SVM algorithm[1] follows the basic framework of classification as shown below:

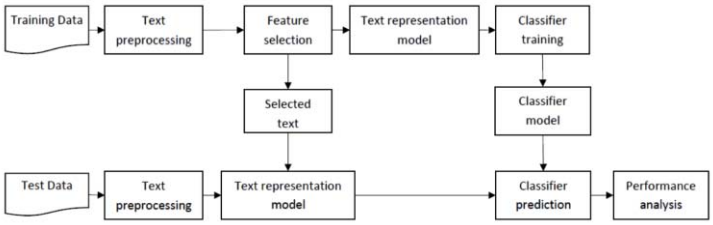


Fig. 1. General architecture for classification

This survey paper summarizes, discusses and compares the recent work of researchers, journal authors and data scientists working on novel ideas of classifying video content published on the web based on their genres. The work discussed in this survey paper is done in the last 3-4 years. This survey paper consists of 6 sections which are presented as follow. Section 2 discusses the background: presented various techniques used by different authors on the classifying problem of video content. Section 3 consists of extensive discussion on existing methodologies. Firstly the technique of bi directional LSTM is discussed, which uses plots of the video to classify its genre. Secondly, the method of Deep Neural Network (DNN) has been discussed for performing classification based on the poster of a movie video followed by a brief discussion on CNN layers classification technique. Section 4 lists the information about famous datasets that are publicly available. Section 5 summarizes the various methods categorized by us.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:40:37 UTC from IEEE Xplore. Restrictions apply.

Section 6 discusses the future scope and challenges in this domain and concludes this survey paper.

II. BACKGROUND

Research papers by different authors can be classified on the basis of approaches they have adopted.

Samuel S. & Chokshi [2] together dissected the film classification by utilizing banners of motion pictures by a profound learning model that distinguishes the film sort passed on by a film banner. To track down an ideal model, they executed modified variants of three standard profound learning structures for picture order: ResNet-50, VGG-16, and DenseNet-169. DenseNet-169 gave us the best outcomes, beating the other two models and models planned in past endeavours at the problem. They expressed that, the utilization of an item identifier calculation like YOLO, assortment of better information, and use of other convolutional neural organization (CNN) designs to the issue.

On comparable lines Gabriel and Kris [3] effectively portrayed that ResNet34 and a custom engineering were sensibly fruitful at foreseeing a banner's classifications. This explains that there is some steadiness in movie visual parts remembered for film banners of a given sort.

Yadav, A. and Vishwakarma, D. K [4] investigated a bound together framework of deep networks for type characterization utilizing film trailers. Their work meant to break down the connection between the feelings inspired by the film trailers and how they contribute in settling the multi-mark type grouping problem. The proposed novel system is approved by performing cross-dataset testing for three huge scope datasets, to be specific LMTD-9, MMTF14K, and ML-25M datasets. Extensive analyses showed that the proposed calculation beat all the best in class strategies.

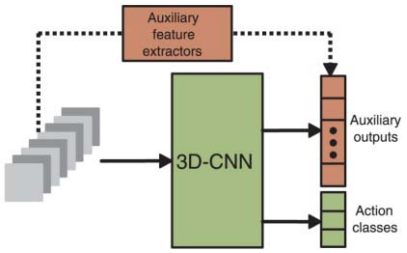


Fig. 2. Explanation of 3D CNN model

Shambharkar, P. G., and Doja, M. N. [5] did the film trailer grouping utilizing deer chasing advancement based profound convolutional neural organization in video sequences. Features was prepared and ordered with the assistance of upgraded deep convolutional neural network (DCNN) for the film trailer classifications. The consequences of the proposed strategy were contrasted and separating measure and without sifting measure in which 95.23% of precision were accomplished for the recommended approach with sifting and 90.91% of exactness was accomplished for the proposed approach without separating measure.

Varshit, Vishal, Rohit, Krishna Reddy, Mamidi [6] anticipated the Genre and Rating of a Movie Based on its Synopsis. They made the Multi-Language Movie Review Dataset (MLMRD). The dataset comprises classification, rating, and summary of a film across various dialects, to be specific Hindi, Telugu, Tamil, Malayalam, Korean, French, and Japanese. In this work, they give different model designs that can be utilized to anticipate the class and the rating of a film across different dialects present in our dataset dependent on the outline.

K Sivaraman and Gautam Somappa [7] dealt with distinguishing the class of a film by investigating simply the visual highlights of its trailer. In this they made a vision system that can do likewise, precisely. They talked about the methodologies they took and exploratory perceptions. The commitments of this work are: (1) proposes a neural organization (in view of VGG) that can characterize film trailers dependent on their sorts; (2) release a curated dataset, called YouTube-Trailer Dataset, which has more than 800 film trailers crossing more than 4 types, accomplishing a precision of 80.1% with the spatial highlights, and 85% with utilizing LSTM and set these outcomes as the benchmark for this dataset.

Y.F. Huang and S.H. Wang [8] thought of a methodology that utilizes the SAHS (Self Adaptive Harmony Search) calculation in choosing highlights for various film types. After this, a Support Vector Machine is stacked with a bunch of highlights which comprises of 277 highlights separated from each film trailer. After this, they utilized a greater part casting a ballot method to gauge the class of the film. They utilize both visual just as sound qualities in this method and utilize a limit of 25 highlights to recognize two distinct types of motion pictures.

S. Ji, W. Xu, Kai Yu [9] chipped away at paper that thinks about the computerized acknowledgment of human activities in observation recordings. In this paper, they built up a novel 3D CNN model for activity acknowledgment. This model concentrates highlights from both the spatial and the transient measurements by performing 3D convolutions, subsequently catching the movement data encoded in various nearby frames.

III. EXISTING METHODOLOGY

A. Bi-Directional LSTM

One major domain to classify movie genre is by taking plot summaries as a dataset and training different models on them

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:40:37 UTC from IEEE Xplore. Restrictions apply.

TABLE I. COMPARISON OF BENCHMARK TECHNIQUES

Year of publication	Dataset used	Technique	Observation	Accuracy
2020[10]	Approx. 21,000 movie posters dating from 1913 to 2019	Gram layer	Genre classification of video content based on poster classification using Inter-Channel features utilises Gram layers, identifying the characteristics of posters and the correlations among every genre, thus engendering better accuracy multi-genre classification. Gram layer brings better focus on videos based on content features of posters than on the attention mechanism, which focuses on basic highlights of photos.	Fscore-0.466 Accuracy-0.45 Precision-0.57 Recall-0.45
2020 (ICICCS)[11]	Movie trailers from Youtube playlists	3D Convulational Neural Networks	The benefit of 3D CNN is its capacity to catch the spatial just as the fleeting data present in a film trailer. 3D CNN had an accuracy of 82.14% which is more noteworthy than different models which use simply 2D CNNs	Accuracy -82.14%
2020 (ICICT)[12]	Deap and Decaf datasets containing movie clips	EEG, MEG, MCA	Identification based on the feelings that are seen by the watchers while watching media cuts is accomplished. The analysis demonstrated that the coupling of magnetoencephalographic signs and multimedia media signals gives higher performance idicator. The high exactness is acquired with this combination alone is on the grounds that the MEG signals have low commotion, accordingly, giving more spatial granularity of brain exercises	RF-0.722 SVM-0.55 KNN-0.59 NB-0.46
2018 (IEEE)[13]	MovieLens dataset of plot summaries	Bidirectional LSTM (Bi-LSTM).	Preparing Bi-LSTM network subsequent to separating the plot outlines into their sentences and intertwining the forecasts for singular sentences beat preparing the organization with the entire plot synopses with the restricted measure of information.	Precision- 67.75% Recall-67.61% F-score-67.68%
2019 (IEEE)[14]	ImageNet dataset	Image based Deep learning model, CNN, LSTM	In the picture cycle, a progression of key-outline highlights got by utilizing the CNN so the LSTM used to become familiar with the unique highlights. Results demonstrated that the deep learning model displays unrivaled execution than the customary techniques. The model based on Deep learning improves the RR% about 2% than the deep learning model introduced in 2017.	CNN+LSTM-77.8% Proposed method without CNN-78.8% , and with CNN 79.7%
2018 (Springer)	IMDB dataset for plot summaries	Topological Data Analysis, NLP	Various general research models that are widely used such as such as Logistic Regression for binary content characterization, Multinomial Naive Bayes and furthermore can perform better and beat further developed procedures like neural organizations when in multi-label text grouping. On the errand of characterization of videos as indicated by four most regular classes, this acquired a Jaccard score of 54.8%, F-score of 71.88%, and hit pace of 83.3%	F-score 71.88% Jacquard index 54.8%
2018 (Springer)	Wikipedia Movie Plots	Text Mining	This paper proposes a framework which takes film plot to be an information and naturally gives a small amount of types present in that film as a yield. Application of modelling procedures, like EM algorithms and LDA, performance scores are boosted. Insight of group theory as film audit can be utilized as a dataset to anticipate the film type data.	Error 7.69%
2010 (ACM international conference)[15]	Movie Trailers	Shot Boundary Detection, Scene Categorization	Genre Classification by means of Scene Categorization utilizes a methodology wherein every trailer is decayed into an assortment of key frames by the means of the model i.e. shot boundary analysis. From these key frames, we use best in class scene indicators and descriptors to extricate highlights. This permits to address trailers applying a bow model with shot classes as vocabularies. The genre grouping task was drawn closer by planning bow transiently organized trailer highlights to four significant level film sorts: action, comedy, drama or thrillers	GIST-71.8% CENTRIST-74.7% W-CENTRIST-73.2%
2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)[16]	Small dataset of movie clip samples	Gaussian kernel SVM prediction model with VC dimension	While implementing support vector machine (SVM) arrangement model and a logistic regression (LR) model was set up to extricate highlights from test information and their test mistake in-example was compared. The accuracy of the calculation expanded to 92% with a more modest VC measurement and less over fitting.	Prediction accuracy of around 92% and 91% were obtained

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:40:37 UTC from IEEE Xplore. Restrictions apply.

Augmented Human Research (2020)[17]	30,000 images From IMDB using a web crawler	Neural networks with knowledge transfer learning	A methodology for the age of a multi-label film sort forecast from its banner utilizing neural organizations are used in the paper. The methodology chips away at two fronts; one is pointed toward making an enormous, various and adjusted dataset for film type forecast. The subsequent front includes re-examining the issue to more straightforward single-name multi-class characterization and producing a multi-label forecast on a given film banner as input.	
2020[18]	500,000 subtitles in Dutch and English for a broad variety of movies.	AdaBoost, C4.5, Naive Bayes, Random Forest, and SVM	Text-based video genre classification concluded that Random Forest and SVM were found to be the strongest algorithms amongst all. Both algorithms are known to be effective on larger datasets. The experimental results showed that of the three feature types, the content-specific features result in the most accurate classifier.	F1-score Linear SVM-0.88 Random Forest-0.83
ASC 2017[19]	LMTD dataset for movie trailers	Multi-Label approach based on Convolutions through Time, multi-label classification	A new model based on neural architecture is proposed relying on convolutional neural networks to perform multilabel video genre classification. It even extends itself to compare this approach with the current state-of-the-art methods for video genre classification. The proposed architecture's performance outplays all the compared models and architectures by the means of various comparison techniques and hence makes the classification even more efficient.	
2019 (ICIT)[1]	Public Film Database	TF-IDF and SVM	It centers around the different advances that should be finished during the process of classification, for example, text word division, including designing, text portrayal, and so on, and plans and executes a film grouping framework dependent on SVM calculation.	

to get results. Several papers are already published on this area of work. Ali Mert Ertugrul et al [13] proposed a classification model called the bi directional LSTM model which showed promising results and accuracy. Plot summaries which are taken as the primary dataset consists of various sentences and assuming that these summaries reflect genre of movie are separated into different sentences. Each sentence is trained with this model and genre of each sentence is generated. In this model dataset as movie names is retrieved from MovieLens[20]. Movie plot summaries are sent in as input and genres are retrieved as the output. This process is done through OMDB API[21]. After data is collected it is pre-processed and represented as semantic and systematic relation with the words.as the length of each lot summaries is different, using data embedding and skip-gram model each input is presented in a vector representation having a dimension of 300.

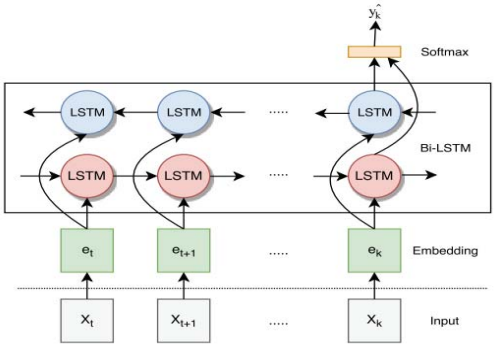


Fig. 3. Bi-Directional LSTM model

BI-LSTM is basically an RNN model which improves and learns complex interdependencies with passage of time. It basically has two neural networks. These are: the forward neural layer which handles and models the previous/preceding text and the other is backward neural network which models the upcoming/following context of plot summaries. These layers are shown in this fig.

Another interesting technique used to classify movie genre based on text summaries is proposed by Quan Hoang et al [22]. In this paper, in contrast to previous techniques discussed, pure, generic machine learning models were implemented as: Naive Bayes, Word2Vec+XGBoost and Recurrent Neural Networks (RNN).This paper discussed these techniques with taking dataset of around 250,000 movie videos and eve achieved f-score of 0.5 with accuracy of around 80 percent.

B. Deep Neural Networks (DNN)

In this section we discuss some work done in classifying movie videos based on genres through Poster images. Various people [10][23][14]have taken up this approach and implemented different models to improve accuracy and efficiency. Classification of genres with Deep neural network is state -of-the-art work done in this field by Wei-Ta Chu [23] et al.

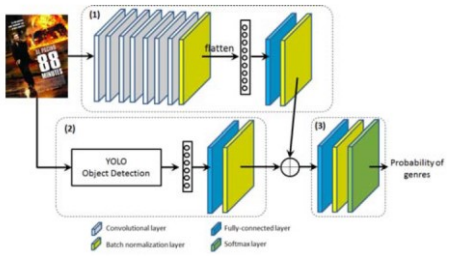


Fig. 4. DNN model using YOLO algorithm

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:40:37 UTC from IEEE Xplore. Restrictions apply.

In this paper authors discuss classification based on DNN model along with object identifier algorithm YOLO[24] as depicted in Fig 3. An overview may explain that large dataset was taken into account along with the metadata of movie posters from IMDB website while crawling across western movies from the era of '80s to 2015. A fixed resolution ranging from 89 × 132 to 300 × 581 was taken into account for poster images. DNN model makes a training set of all the images as D = {A,B }, where A = (a1, a2, ..., aN) is the set of n poster images and B0 = (b1 ,b2 , ...,bN) is the corresponding genre information. The vector ci = (ci'1, ci'2, ..., ci'M) is a binary vector, where ci,j = 1 shows what genre is connected with which of its corresponding poster image.

One problem that is generally encountered in classifying genre through poster images is that a single image may portray more than one genre, which is also called the multi label image genre classification problem.

Apart from DNN Samuel Sung et al. [2] worked on movie genre classification with poster images using other deep learning models such as: ResNet-50, VGG-16, and DenseNet-169. DenseNet-169.and in conclusion DenseNet-169 was the best performing model as it achieved the highest F1 score amongst others.

C. Convolutional Neural Networks (CNN)

Convolutional Neural Network is a very unique and special tool used for picture classification which can be also extended to movie frames classification and is widely used by various tech giants such as Facebook and Amazon for segregating images for their clients.

In CNN the main task is to accept an image as an image and create or define its class. A multidimensional vector is created every time an image is taken as an argument by the neural layer. For instance if a resolution of an image is 250X 250, then the dimension of the vector created will be of 250X250X3 (3 as RGB values).

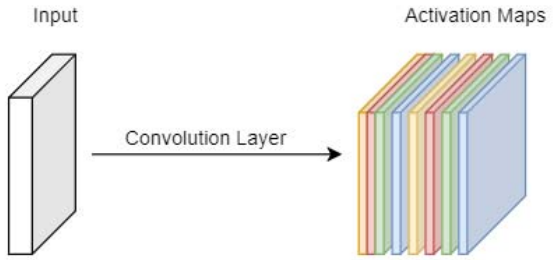


Fig. 5. 3D CNN model

Based on CNN various papers have been already published which attempts to classify movie genres. First paper we discuss here is by Gabriel S. Simoes et al [25]. Borrowing ideas from image processing concepts, this paper took a dataset of around 3500 known movie videos. Subsequently in the latter half of the paper the CNN model produced results are compared with feature extraction models such as GIST CENTRIST and W CENTRIST.

These feature extraction models which are used in various papers are summed up as follows:

- GIST: The GIST model gives a singular, comprehensive feature descriptor for the input image that encodes semantic data illustrating features of the image scene. The semantic attributes of interest illustrates high level properties of the scene, like its roughness, ruggedness, and openness.

- CENTRIST: The CENsus TRansform hISTogram, is a visual feature extracted created for recognition of the semantic category of original scenes and indoor atmosphere, e.g. coasts, forests, bedrooms, streets, living rooms, etc. They demonstrated that CENTRIST creates outstanding results.

- W-CENTRIST: Both GIST and CENTRIST models deny hues indicator on creating descriptors. Therefore, W-CENTRIST is a model of the CENTRIST feature extract which collects colour data in addition to intensity stats.

Furthermore, In [15] a file of around 144 new trailers which were not the part of dataset used for training the data was tested on the same model of scene categorization. With, original parameter settings model was tested and an accuracy of around 72% was measured which more than the accuracy was obtained on the training set.

In [25] the authors also introduced a novel method In this paper supported by a CNN architecture to classify movie trailers according to their genres, namely CNN-Motion (Convolutional Neural Networks for Movie Trailer Classification). Finally the accuracy with this model comes out to be around 74% with this mode.

IV. DATASET

A variety of summaries, plots, posters, trailers are used in this research to depict the movie genre precisely in the distinct categories of 'romance, action, comedy, horror' etc. Dataset used in this research topics differs with the different approaches used in predicting the movie genre, may it be through the features of movie trailers, movie posters or using the descriptive movie plots which tell the do and about of the movies.

- For the movie trailers, the data has been extracted through the playlist provided in Youtube which gives us over 1000+ trailers for each category.
- For plots, the main dataset used is the IMDB Film Review Dataset which contains over 50,000 movie reviews of all the wide variety of genres.
- For posters, the dataset is either extracted from the IMDB website or using the Full MoiveLens dataset from Kaggle consisting of data from TMDB and groupLens for over 45,466 movies which have the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:40:37 UTC from IEEE Xplore. Restrictions apply.

data of their genres and posters of 500 x 750 resolutions in size which perfectly fits.

V. CONCLUSION AND FUTURE WORK

Today we live in a world where having access to numerous video based content is not only easily accessible but also abundant. This video content can vary in genres and forms. Analysing them and operating different techniques, models and algorithms helps to put large data into separate 'buckets' of genres so that it is easier for the user to segregate and chose a video based on his/her choice.

In our paper, we conducted a thorough survey of the latest papers in the field of classification of videos and analysed state of the art techniques and methods applied by different researchers and scientists to solve this problem of genre classification. We also categorized these papers into four subcategories on the basis of type of parameter a model takes up to classify for instance: Plot of the video, poster of the video, audio and video features etc. Apart from algorithms and methods, we also analyse the popular and publicly available datasets and list them in our paper.

There has been a rise in the research for new state of the art classifications in the last few years and it is only going to increase in the future. There have been lots of opportunities and challenges in the problem of making visual content easily accessible to the users considering that there is plethora of options available to consume. In the last few years a lot of research has been conducted in classification models and techniques and it will be interesting to explore the intersection of genre classification models and classification of videos based on other parameters such as adult content or false/rumor content, and also classification based on type of music used in a movie, grouping dataset based on message a video is trying to portray and many more. Basically having more parameters to classify video contents opens up lot of space to explore in future .There are also a lot of challenges that need to be tackled such as scalability on real world networks of content, along with numerous opportunities and areas to be explored while going in this direction of area of work.

References

[1] N. Fei, "Movie genre classi fi cation using TF-IDF and SVM," pp. 3–8.

[2] R. Chokshi and S. Sung, "Classification of Movie Posters to Movie Genres," p. 2020, 2020.

[3] G. Barney and K. Kaya, "Predicting Genre from Movie Posters," *Semant. Sch.*, 2019.

[4] A. Yadav and D. K. Vishwakarma, "A unified framework of deep networks for genre classification using movie trailer," *Appl. Soft Comput. J.*, vol. 96, p. 106624, 2020, doi: 10.1016/j.asoc.2020.106624.

[5] P. G. Shambharkar and M. N. Doja, "Movie trailer classification using deer hunting optimization based deep convolutional neural network in video sequences," *Multimed. Tools Appl.*, vol. 79, no. 29–30, pp. 21197–21222, 2020, doi: 10.1007/s11042-020-08922-6.

[6] V. Battu, V. Batchu, R. R. Reddy, M. K. Reddy, and R. Mamidi, "Predicting the genre and rating of a movie based on its synopsis," *Proc. 32nd Pacific Asia Conf. Lang. Inf. Comput. PACLIC 2018*, no. December, pp. 52–62, 2018.

[7] S. K. S, G. Somappa, K. S. Sivaraman, and G. Somappa, "MovieScope : Movie trailer classification using Deep Neural Networks," pp. 2–8, [Online]. Available: https://virginia.app.box.com/s/hhe7xeq96d99yuqn1nr1q9yxdhs7ey9r.

[8] Y. F. Huang and S. H. Wang, "Movie genre classification using SVM with audio and video features," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7669 LNCS, no. December, pp. 1–10, 2012, doi: 10.1007/978-3-642-35236-2_1.

[9] S. Ji, W. Xu, M. Yang, and K. Yu, "3D Convolutional neural networks for human action recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 221–231, 2013, doi: 10.1109/TPAMI.2012.59.

[10] J. A. Wi, S. Jang, and Y. Kim, "Poster-Based Multiple Movie Genre Classification Using Inter-Channel Features," *IEEE Access*, vol. 8, pp. 66615–66624, 2020, doi: 10.1109/ACCESS.2020.2986055.

[11] P. G. Shambharkar, P. Thakur, S. Imadoddin, S. Chauhan, and M. N. Doja, "Genre Classification of Movie Trailers using 3D Convolutional Neural Networks," *Proc. Int. Conf. Intell. Comput. Control Syst. ICICCS 2020*, no. Iciccs, pp. 850–858, 2020, doi: 10.1109/ICICCS48265.2020.9121148.

[12] H. Pothina, G. Gopakumar, and A. Narayanan, "Multimedia Genre Prediction by Analysing Neurophysiological Signals," *Proc. 5th Int. Conf. Inven. Comput. Technol. ICICT 2020*, pp. 364–368, 2020, doi: 10.1109/ICICT48043.2020.9112475.

[13] A. M. Ertugrul and P. Karagoz, "Movie Genre Classification from Plot Summaries Using Bidirectional LSTM," *Proc. - 12th IEEE Int. Conf. Semant. Comput. ICSC 2018*, vol. 2018-Janua, pp. 248–251, 2018, doi: 10.1109/ICSC.2018.00043.

[14] C. H. Chou and P. C. Jen, "Image based deep learning model for movie trailer genre classification," *2019 IEEE 8th Glob. Conf. Consum. Electron. GCCE 2019*, pp. 1064–1066, 2019, doi: 10.1109/GCCE46687.2019.9015293.

[15] H. Zhou, T. Hermans, A. V. Karandikar, and J. M. Rehg, "Movie genre classification via scene categorization," *MM'10 - Proc. ACM Multimed. 2010 Int. Conf.*, pp. 747–750, 2010, doi: 10.1145/1873951.1874068.

[16] H. Wang and H. Zhang, "Movie genre preference prediction using machine learning for customer-based information," *2018 IEEE 8th Annu. Comput. Commun. Work. Conf. CCWC 2018*, vol. 2018-Janua, pp. 110–116, 2018, doi: 10.1109/CCWC.2018.8301647.

[17] K. Kundalia, Y. Patel, and M. Shah, "Multi-label Movie Genre Detection from a Movie Poster Using Knowledge Transfer Learning," *Augment. Hum. Res.*, vol. 5, no. 1, 2020, doi: 10.1007/s41133-019-0029-y.

[18] A. Pal, A. Barigidad, and A. Mustafi, "Identifying movie genre compositions using neural networks and introducing GenRec-a recommender system based on audience genre perception," *Proc. 2020 Int. Conf. Comput. Commun. Secur. ICCCS 2020*, 2020, doi: 10.1109/ICCCS49678.2020.9276893.

[19] J. Wehrmann and R. C. Barros, "Movie genre classification: A multi-label approach based on convolutions through time," *Appl. Soft Comput. J.*, vol. 61, pp. 973–982, 2017, doi: 10.1016/j.asoc.2017.08.029.

[20] "https://grouplens.org/datasets/movielens/." .

[21] "http://www.omdbapi.com/." .

[22] Q. Hoang, "Predicting Movie Genres Based on Plot Summaries," *arXiv*, pp. 1–12, 2018.

[23] W. T. Chu and H. J. Guo, "Movie genre classification based on poster images with deep neural networks," *MUSA2 2017 - Proc. Work. Multimodal Underst. Soc. Affect. Subj. Attrib. co-located with MM 2017*, pp. 39–45, 2017, doi: 10.1145/3132515.3132516.

[24] "https://appsilon.com/object-detection-yolo-algorithm/." .

[25] G. S. Simões, J. Wehrmann, R. C. Barros, and D. D. Ruiz, "Movie genre classification with Convolutional Neural Networks," *Proc. Int. Jt. Conf. Neural Networks*, vol. 2016-Octob, pp. 259–266, 2016, doi: 10.1109/IJCNN.2016.7727207.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:40:37 UTC from IEEE Xplore. Restrictions apply.

Multibeam antenna using cavity-based waveguide structure with extended hemispherical lens at millimetre wave

Shakti Singh Chauhan,[1]✉ Harikesh Dalal,[2] Mahesh P. Abegaonkar,[1] and Ananjan Basu[1]

[1]*Centre for Applied Research in Electronics, Indian Institute of Technology Delhi, New Delhi, India*

[2]*ECE Department, Delhi Technological University, Delhi, India*

✉Email: shakti.uit@gmail.com

A multibeam antenna with a single input port using a cavity-based waveguide structure is implemented. The proposed structure uses a cylindrical cavity and a higher-order TM_{21} mode is excited in it. The higher-order mode is to divide the incoming power in three different directions. The beam in three different directions is confined by using a 3D-printed extended hemispherical lens integrated at three openings at the cylindrical cavity. The proposed structure has measured an impedance bandwidth of 1.7 GHz from 56.9 to 58.6 GHz. The achieved beam directions for the proposed antenna structure is 0, 87 and 270°. The measured gain of the antenna for each beam directions is better than 12 dBi.

Introduction: Wireless Body Area Networks (WBAN) has attracted much attention recently. The frequency band at 60 GHz is one of the promising candidates that can be used for WBAN. Antennas have been proposed for off-body and on-body links to be used when placed over the body at V band [1–3], with less focus for the antennas used in external nodes to set up a link with a node over the body. The omni-directional antennas can be used as an external node to set up off-body links, but they suffer from low gain. So a high-gain multibeam antenna is proposed to be used in the external node. The high gain will help to combat the losses faced by the wireless link, and multibeam antennas will help to improve the coverage by radiating beams in different directions. The high gain is achieved either by using an array or by using a dielectric lens over the radiator. The dielectric lens has an upper hand over the antenna array as for arrays number of antenna elements needed is large with the complex feeding mechanism. The dielectric lens will focus the electromagnetic wave and thus increase the gain of the antenna structure [4]. The dielectric lens antenna can also provide beam steering by switching the radiating elements beneath it [5]. The beam steering can also be done using phased array systems, but the phased array will also have a complex geometry. Both phased arrays and dielectric lens configuration will require a large number of radiating elements for beam steering. Thus, a multibeam antenna configuration is proposed to be used, which will be a much simpler approach to enhance coverage in different directions. The narrow beams will naturally leave regions without coverage, but in practical cases, the antenna can usually be oriented such that coverage is not required there (e.g. regions covered by furniture in a hospital room).

The proposed structure is a waveguide-based structure, which is used to excite a higher-order cylindrical cavity. The cylindrical cavity is used to bifurcate the input incoming power into three directions which are 0°, 90°, 270° angles, which is further confined by using a reduced extended hemispherical lens. The proposed antenna provides high gain with multibeam radiation characteristics. The multibeam antenna helps in the angular coverage in three orthogonal directions. The rest of the paper is organised as follows: The antenna design is discussed in the Section 2 and is followed by the results and conclusion in Sections 3 and 4, respectively.

Antenna design: The proposed structure is based on exciting a higher-order mode in a cylindrical waveguide cavity. The modes of the cavity depend on the radius of the cavity structure. The whole idea is to excite a higher-order TM_{21} mode in the cylindrical cavity. Thus, the radius of the cavity is decided by the mode excited. The eigenmode analysis is used to determine the radius of the cavity to excite desired higher-order mode at a frequency of interest. The height of the cavity is kept at 2.4 mm

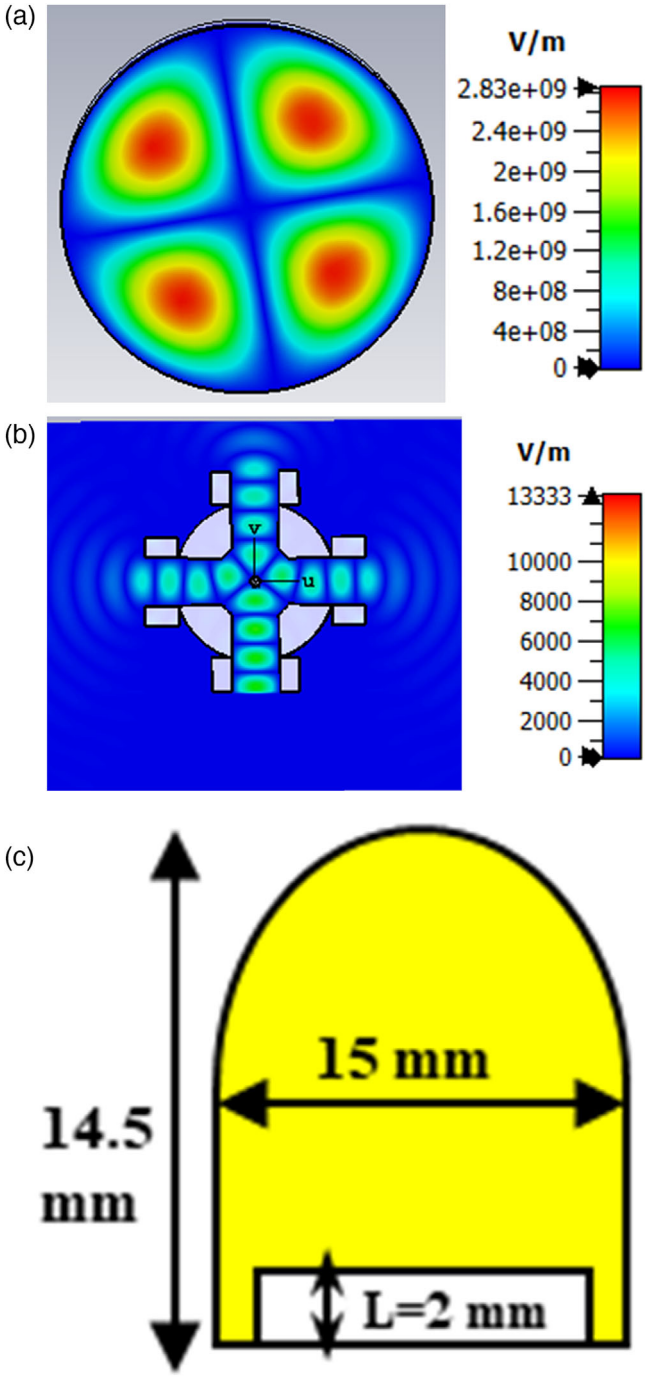


Fig 1 *(a) E-field distribution of the cylindrical cavity in eigenmode, (b) E-field distribution of the antenna structure with a metallic post at 61 GHz, (c) dimensional view of the 3D printed lens structure*

according to a shorter length of the WR 19 waveguide, which is used to feed the cylindrical cavity. The electric field distribution in the cylindrical metallic cavity with a radius of 4 mm in eigenmode solver is shown in Figure 1(a). The TM_{21} mode will exist at around 61.26 GHz. The waveguide is used to excite the cylindrical cavity, and three openings are made so that the antenna will radiate in three different directions. A cylindrical post of 1 mm diameter is made at the centre of the cylindrical structure. The higher TM_{21} mode excited in the cylindrical structure bifurcates the incoming wave into three different directions as shown in Figure 1(b). With the help of these opening, whose cross-sectional dimensions are kept the same as WR 19 dimensions (4.8*2.4 mm), the structure will start radiating in three different directions. Thus, a multibeam single input antenna is implemented. The broad beam radiation is seen in three orthogonal directions. In order to confine beams, a 3D-printed extended hemispherical lens is integrated over the waveguide structure. The extended hemispherical lens is made using polylactic acid

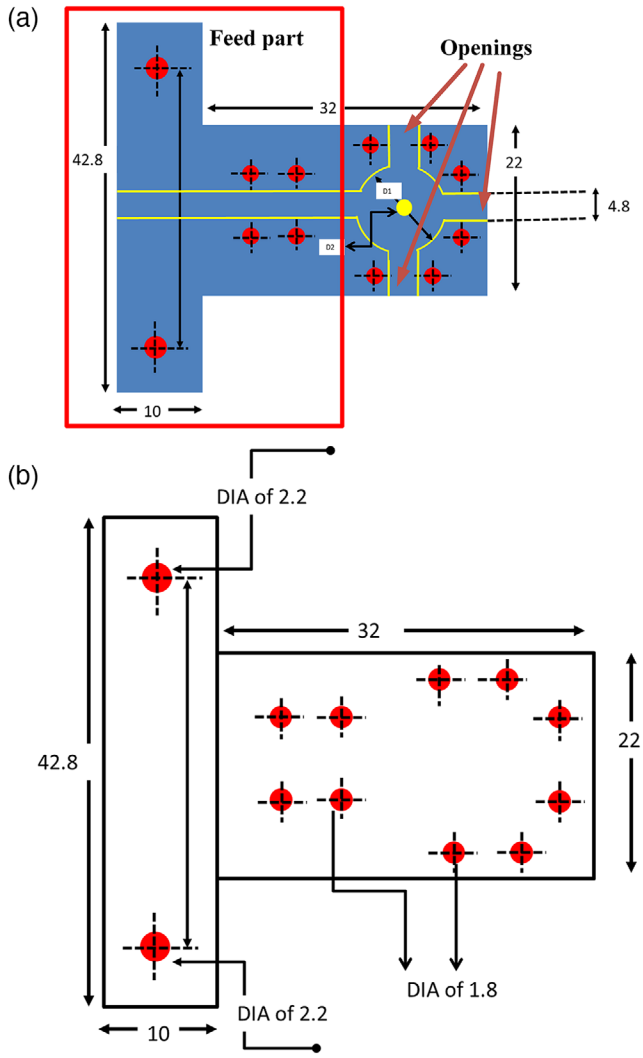


Fig 2 *Dimensional view of the structure (a) top view of the bottom piece, (b) top view of the top piece*

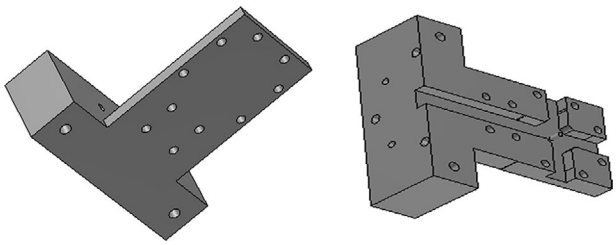


Fig 3 *3D view of top and bottom piece of the antenna structure*

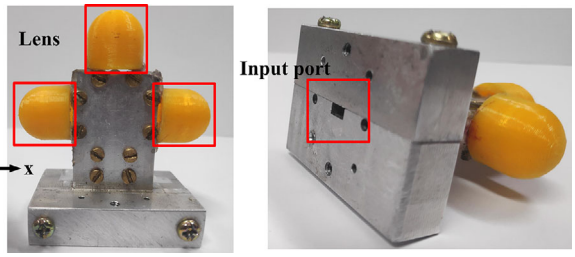


Fig 4 *Fabricated final antenna structure with lens integration*

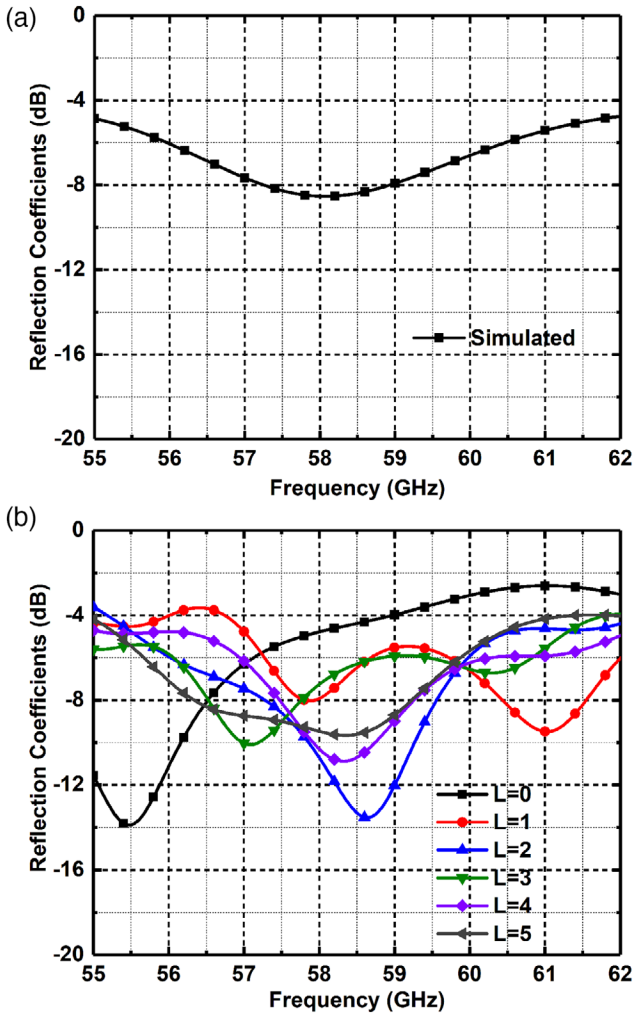


Fig 5 *Simulated reflection coefficients (a) antenna without lens integration, (b) antenna after lens integration with a parametric variation of the length of an air pocket in the lens*

(PLA) based on the fused deposition modelling (FDM) technique. The dielectric properties of PLA have been given in [6]. The dimensions of the lens are calculated from the equations given in [7, 8] and are shown in Figure 1(c). The air pocket is made at the bottom of the lens so as to do impedance matching. The cylindrical cavity-based waveguide structure is fabricated in two parts using mechanical machining on aluminium. The top and bottom parts are shown in Figure 2, where D1 = 8 mm, diameter of the cylindrical structure and D2 = 1 mm, the diameter of the cylindrical post. The 3D structure in different perspective planes is shown in Figure 3. The two parts are fitted together and the lens is integrated over three openings as shown in the final proposed structure in Figure 4. Thus, the lens-integrated multibeam antenna structure is proposed. In the next section, the simulated and the measured result of the proposed antenna structure is discussed.

Results: The s-parameters of the proposed structure without lens are shown in Figure 5(a). The resonance occurs at a frequency 58 GHz with a reflection coefficient less than –10 dB. As the lens integrated over three openings of the structure, the change in the reflection coefficient is clearly seen in Figure 5(b) for the case of L = 0 mm, where L is the length of air cavity made in 3D-printed lens. In the lens structure, a small cylindrical air pocket is made and its length is varied from 0 to 5 mm with a step size of 1 mm. The variation in the length of the air pocket helps in impedance matching. From Figure 5(b), the length is selected to be L = 2 mm for the final proposed structure. The comparison between the measured and the simulated reflection coefficient is shown in Figure 6(a). The measured impedance bandwidth is 1.7 GHz from 56.9 to 58.6 GHz. The deviation observed in the reflection coefficient may be due to fabrication tolerances, still the measured results are in agreement with the simulated result. The E-field distribution is shown in Figure 6(b), and it is clearly seen that the incoming wave is bifurcated in three different directions, which further illuminates the

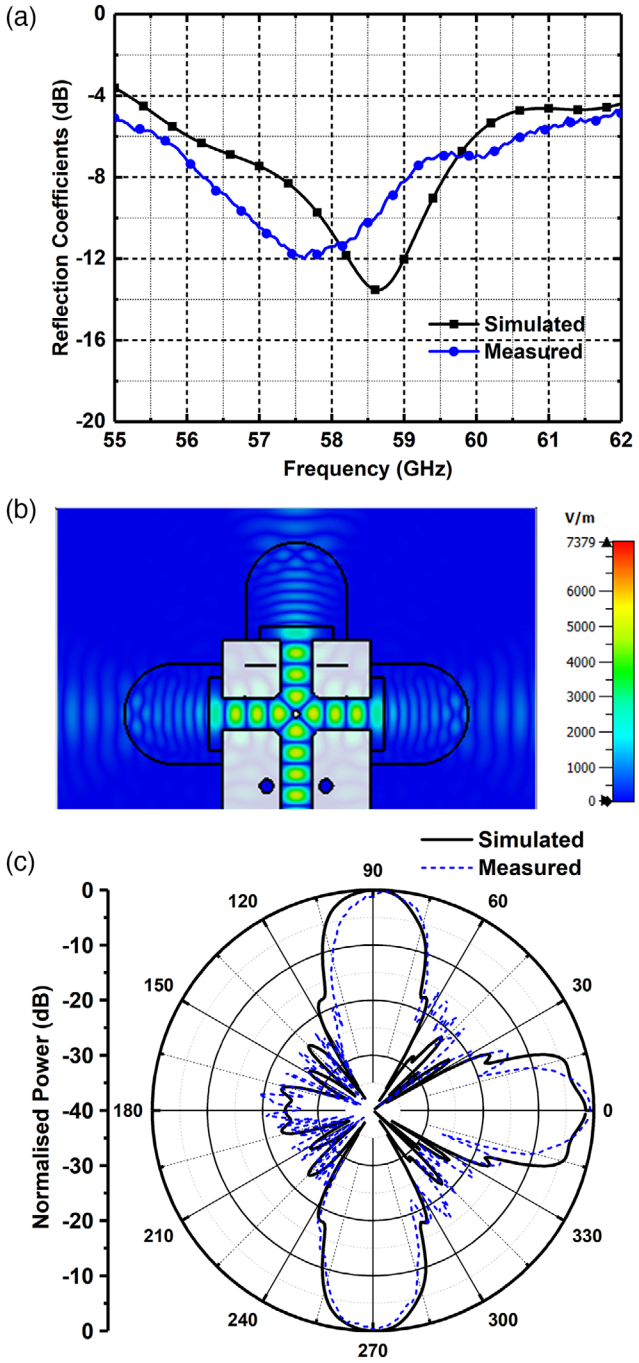


Fig 6 *Comparison between the simulated and measured results of the final fabricated antenna structure (a) reflection coefficients, (b) E-field distribution at 57GHz, (c) 2D normalised xy plane (H-plane) pattern at 57 GHz*

integrated lens structure. The normalised radiation pattern of the proposed structure is measured in xy-plane (H-plane) and compared with simulated results in Figure 6(c) at 57 GHz. The measured beam directions for beam1, beam2 and beam3 of the structure are 0°, 87° and 270°, respectively at 57 GHz. The measured 3 dB beamwidth of the antenna for beam1, beam2 and beam3 are 17°, 19° and 18° respectively, at 57 GHz. The simulated radiation efficiency of the antenna is 66.5% at 57 GHz. It is clearly seen from the normalised radiation pattern that a singly fed antenna with multibeam pattern is achieved in the proposed design. The measured gain of the antenna at 0°, 87° and 270° are 12.2, 12.8 and 12.7 dBi, respectively.

Conclusion: A multibeam antenna using a cavity-based waveguide structure is implemented. The proposed structure uses a cylindrical cavity, and a higher-order TM_{21} mode is excited and is used to divide the incoming power in three different directions. The beam in three different directions is confined by using a 3D-printed extended hemispherical lens integrated at three openings in the cylindrical cavity. The measured bandwidth is 1.7 GHz and a gain better than 12 dBi is achieved for all the three beam directions.

© 2021 The Authors. *Electronics Letters* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
Received: *22 January 2021* Accepted: *1 March 2021*
doi: 10.1049/ell2.12154

References

1 Chahat, N., et al.: Wearable endfire textile antenna for on-body communications at 60 GHz. *IEEE Antennas Wirel. Propag. Lett.* **11**, 799–802 (2012)
2 Chahat, N., Zhadobov, M., Sauleau, R.: Wearable textile patch antenna for BAN At 60 GHz. In: 2013 7th European Conference on Antennas and Propagation (EuCAP), Gothenburg, Sweden, pp. 217–219 (2013)
3 Puskely, J., et al.: Wearable disc-like antenna for body-centric communications at 61 GHz. *IEEE Antennas Wirel. Propag. Lett.* **14**, 1490–1493 (2015)
4 Bisognin, A., et al.: 3D printed plastic 60 GHz lens: Enabling innovative millimeter wave antenna solution and system. In: 2014 IEEE MTT-S International Microwave Symposium (IMS2014), Tampa, FL, USA, pp. 1–4 (2014)
5 Artemenko, A., et al.: Beam steerable quartz integrated lens antenna for 60 GHz frequency band. In: Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP), Rome, Italy, pp. 758–762 (2011)
6 Felício, J.M., Fernandes, C.A., Costa, J.R.: Complex permittivity and anisotropy measurement of 3D-printed PLA at microwaves and millimeter-waves. In: 2016 22nd International Conference on Applied Electromagnetics and Communications (ICECOM), Dubrovnik, Croatia, pp. 1–6 (2016)
7 Filipovic, D.F., Gearhart, S.S., Rebeiz, G.M.: Double-slot antennas on extended hemispherical and elliptical silicon dielectric lenses. *IEEE Trans. Microwave Theory Tech.* **41**(10), 1738–1749 (1993)
8 Godi, G., Sauleau, R., Thouroude, D.: Performance of reduced size substrate lens antennas for Millimeter-wave communications. *IEEE Trans. Antennas Propag.* **53**(4), 1278–1286 (2005)

Proceedings of the Fifth International Conference on Trends in Electronics and Informatics (ICOEI).
IEEE Xplore Part Number:CFP21J32-ART; ISBN:978-1-6654-1571-2

Multiplex Regulation System With Personalised Recommendation Using ML

Anshul Shroff
Department of CSE
Delhi Technological University
Delhi. India
anshul.shroff223@gmail.com

Bickey Kumar Shah
Department of CSE
Delhi Technological University
Delhi. India
Bickeyshsh721@gmail.com

Aayush Jha
Department of Civil Engineering
Delhi Technological University
Delhi. India
aayushjha452@gmail.com

Amar Kumar Jaiswal
Department of CSE
Delhi Technological University
Delhi. India
jaiswalamar22@gmail.com

Pooja Sapra
Department of CSE
Delhi Technological University
Delhi. India
dspooja17@gmail.com

Manoj Kumar
Department of CSE
Delhi Technological University
Delhi. India
mkumarg@dce.ac.in

2021 5th International Conference on Trends in Electronics and Informatics (ICOEI) | 978-1-6654-1571-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICOEI51242.2021.9453005

***Abstract*—The main aim of this research work is to conduct a feasibility study and requirement analysis for a Multiplex Management system (MMS). This research work includes the design diagrams that facilitate the production of the software and its use in making a frontend and backend for the system according to the elicited requirements, by integrating them. Its alpha version, which strictly adhering to relevant software engineering practices that test the software with suitable methods has also developed. It automates various functions of Multiplex Theater with particular emphasis on ticket reservation, show screening, personnel management, statistical report generation, and increasing ticket sales by giving personalized movie suggestions by learning from the previous booking history or characteristics of users. MMS is a Web application that provides different interfaces for different stakeholders like users, admins, and employees. The admins can add movies and schedule screenings and evaluate the booking statistics and the users and employees can book and cancel tickets. The system also analyzes various factors using Machine Learning to suggest movies to the User to maximize their chances of booking them. The data and model generated from this system are also a precious commodity that can be leveraged successfully to bring in large gains for the Service.**

Keywords—requirement analysis, Machine learning, web application, central Database, Graphical User Interface.

I. INTRODUCTION

The majority of tickets sold are through an online platform compared to the counter at the Multiplex and since a large number of them are not part of big Multiplex Chains developing and maintaining their system is very costly [3]. Also, the data generated through this system is highly valuable from which the Service can derive important analytics and predictions. Multiplexes also employ many people and having some tool to manage them simplifies a lot of work. Manual management of the functions of Multiplex is a cumbersome task because there are multiple screens and simultaneous screenings and there is a scope of various errors. MMS is capable of managing the daily activities of the Multiplex, handle reservations of tickets and also 4 manage employees. MMS tries to automate much of this work and also maximize the revenue in different ways. It increases the efficiency of the employee and also automates the show screening. It helps in giving the movie suggestions and provides the statistical data to the admins and makes informed decisions. This research work includes the analysis of different requirements and all of them are deeply classified into functional and non-functional requirements [1].

A. Functional Requirements

1) Multiplex

- Automate their current multiplex management and ticket reservation system
- Account for employees and staff
- Model the multiplex structure and manage movie and screen timings
- Collecting user data on movie preferences for connecting with their local audience
- Connecting with their users through notifications and updates as well as promotions and branding

2) Customers

- Visually pleasing and intuitive UI design
- Recommend movies and shows similar to their preferences
- Updation of new shows and nearest multiplexes
- Advertising and Promotions

B. NON-functional Requirements

The software needs to be:

- Scalable - As the software needs to be accountable and bring in more multiplexes and more users into their

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:03:02 UTC from IEEE Xplore. Restrictions apply.

fold. The software needs to ensure that it can bear heavier usage and still run as needed.

- Maintainable - The Software needs to be developed such that it is possible for regular maintenance. Since the requirements wouldn't change much once established, regular maintenance would still be required to keep the system performing [2].
- Performance - the system should be able to handle huge loads of requests and the servers shouldn't fail to service the customers during peak times. The database and ticket reservation system should be error-proof and robust.
- Security - As the application collects user data as well as enables online payment systems there must be no security faults and holes in the software that can be exploited by a third party. Security should be of utmost priority and software needs to be designed to ensure the protection of data as well as the safety of online transactions[9].

II. FEASIBILITY STUDY

A. Financial Feasibility

The multiplex management system, being a web application, will subsequently have a cost for web hosting. Since the system does not consist of any multimedia transfer, the bandwidth required for the application is very low. Following freeware software standards, there will be no charges to potential customers [5]. Bug fixes and maintaining/updating the codebase would bear an additional cost for the company [18]. The major source of revenue for the company will be the Multiplexes that opt to use the service to automate their ticket reservation system. The Multiplex Management System can be an extremely attractive alternative to current ticketing systems in place. It can greatly boost their reach as well as retain potential customers by adding a host of secondary services, notifying new shows and offers, and personalized recommendations. The Multiplex is sure to benefit from making use of our system [16]. By offering this service to multiple multiplexes and expanding further users can later choose from multiplexes to view their preferred show. From there it's clear that the Multiplex Management System is financially feasible.

B. Technical Feasibility

The MMS is a complete web application. The Technologies and tools to be used for project development will be open source and freely available. The time limit of developing the product and the ease of using these technologies is synchronized and ensures that our progress matches with projected deadlines. There will be the use of already existing open-source software and APIs such as for the Payment portal.

- Git - Code Management
- Spring Boot
- SQL - Database Management

As the MMS is a web application, hosting the web application would be required. The use of open-source software for parallel development will help the developer teams stick to time constraints as well as keep up with the schedule. From there it's clear that the MMS is technically feasible [4].

C. Social/Legal Feasibility

The MMS uses freely available development tools and provides the system as an open-source system. The maintenance cost will be borne by the company. Since this system is a viable alternate to traditional and outdated ticket reservation systems and also encourages customer interaction with the multiplex, it will have a great impact as a Multiplex Management System[8-10].

D. Risk Feasibility

The feasibility of risk can be determined into different categories of size requirements and requirements changes. size of the codebase will be large as there are multiple stakeholders and as well as a customer portal and a portal for multiplex owners. The size of the Database would be moderately large as multiple multiplexes and users as well as their booking information would build a large temporary database as new bookings will be added and already existing data will be used for the recommender system. The requirements are well defined and understood before the implementation phase and since the requirements are general requirements would be changed only if additional functionality is to be added. Effects of the system on the company revenue matters. The system aims to bring together multiplexes under one umbrella.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:03:02 UTC from IEEE Xplore. Restrictions apply.

Moreover, it helps malls and banks to partner and brings coupons and promos, along with the movie recommender system, which will encourage more customers to be frequently in touch with the cinema. This all points to creating continuous and growing revenue with this system in place. The sophistication of end-users matters a lot. MMS is designed while maintaining the complexity at a very low level. The system is designed to be a highly usable system emphasizing Human-Computer Interaction. Usability is highly improved by providing help documents and making the GUI highly interactive and easy to use [7].

III. LITERATURE SURVEY

MMS is an age-old tradition which has been the primary source of entertainment and hobby for many individuals. It has been more than a century since, the theatre has blessed people with not only entertainment or time pass but has also given them idols to follow and in many cases, worship [13].

As this system has existed for so long in our society, it is only fair to believe that it has been through a tremendous amount of change to still be intact as a practice and be a formidable option for entertainment to this date [8]. This experience of visiting various Multiplex to watch a movie has become more and more of a luxury with the progressing time. For instance, booking of your tickets is a click away and that too with the seat of your choice to give you more of an immersive and a personalized experience. The handling of database with the records of customers and movies has been digitalized which has made the working of these theatres or multiplexes smooth and has definitely enhanced the accessibility of these records and made the current digitalized system more reliable than the traditional handwritten method. It has also made the tracking of employee and brand records much simpler and definitely the modern system of working has given these multiplexes a chance to tie up and generate a good amount of revenue from sponsorships and advertising [6-12].

Thus, the influence of technology has not only just been profitable but also advantageous and less tedious for the functioning of these multiplexes.

IV. METHODOLOGY

The MMS will be made using Incremental Process SDLC Model. An incremental build model is a method of software development where the model is designed, implemented, and tested incrementally (a little more is added each time) until the product is finished. The product is defined as finished when it satisfies all of its requirements. This model combines the elements of the waterfall model with the iterative philosophy of prototyping. The requirements are gathered, and specifications are defined as well as possible at the beginning of the project and an SRS document is produced. After Requirements gathering and specification, requirements are then spitted into several different versions starting with version1, in each successive increment, the next version is constructed and then integrated with the current version. Each incremental version is usually developed using an iterative waterfall model of development. As each successive version of the software is constructed and delivered, now the requirements for the next iteration are decided and incorporated in the next version. We plan to use the Parallel Development Model in which, different subsystems are developed at the same time. It can decrease the calendar time needed for the development, i.e TTM (Time to Market) if enough resources are available.

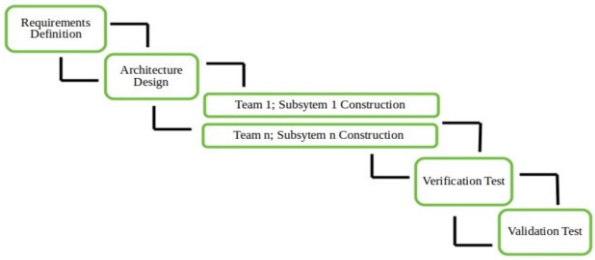


Fig.1 Model Diagram

There are some problems with this model. One is that each new build must be integrated with previous builds and any existing systems. The task of decomposing products into builds not trivial either. If there are too few builds and each build degenerates this turns into the Build And-Fix model. However, if there are too many builds then there is little added utility from each build. The software is then evaluated using some software quality metrics and tested to ensure it is reliable and efficient. In the next step, various documentation is created which are necessary for the maintenance and operation of the software.

A. FACILITIES REQUIRED

As the MMS is a complete web application, emphasis should be given to developing an intuitive and interactive front-end design and an error-free and robust back-end structure. The front end is developed using the popular JS framework while the backend is built along with SQL being used as the primary database. For building the movie recommender system, Machine Learning algorithms will be developed on the Python ML Framework TensorFlow. For code version control and project management Git

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:03:02 UTC from IEEE Xplore. Restrictions apply.

will be the primary tool. Automated Software Testing will be carried out using JUnit.

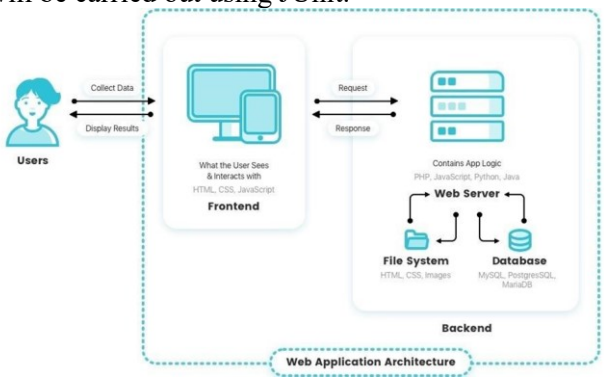


Fig.2 Web application Architecture

In the mentioned fig.2, it is the web application architecture that's shows the relationship and intra connection between the user, front end connection and connection of the frontend with the file system and database. The way and the request and respond is shown in the architecture clearly which is the intermediate od the interface in the frontend and backend steps.

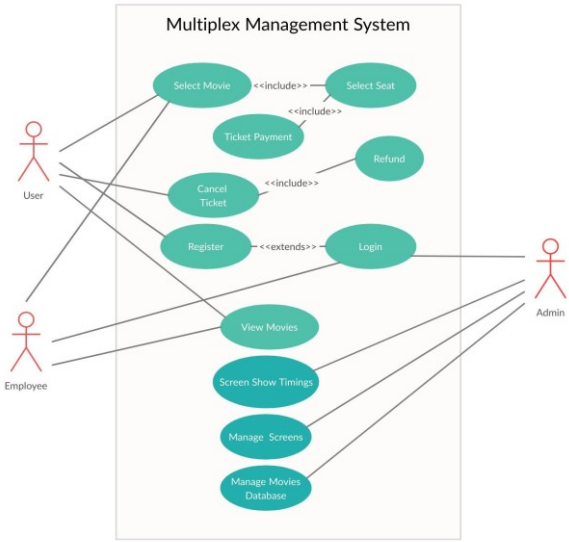


Fig.3 Use case Diagram

The use case diagram is shown in the figure between the user, employee and Admin. It is clearly shown that the user can select movie, cancel tickets and view movies [15]. It can also make the registration process and work. Employee can select movie, login and can view movie. And admin can do 4 tasks of login, screen show timings, manage screens and manage movies database. Intra connection of the tasks are also shown in the pictures.

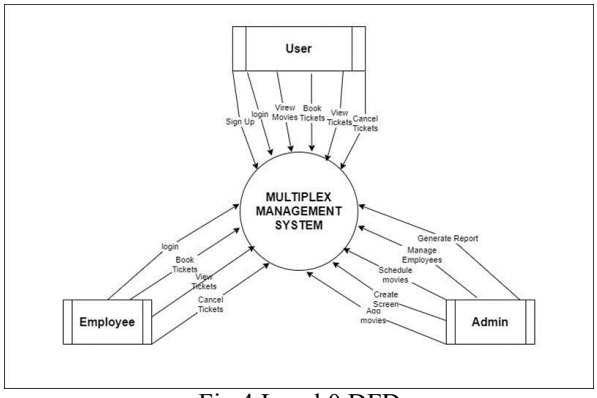


Fig.4 Level 0 DFD

Mentioned in the fig.4 is the data flow diagram of Level 0, It is well shown that how the system is singly related with all the entities. There is a single relationship of all entities with incoming and outgoing arrows.

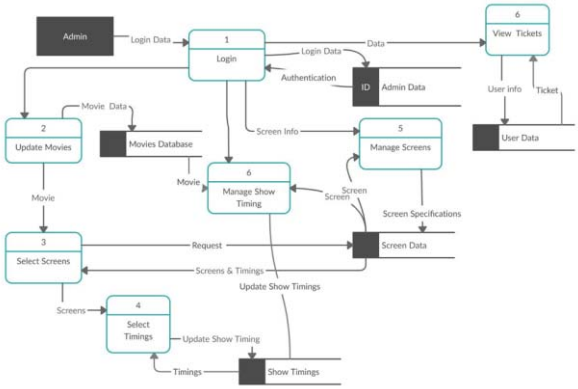


Fig.5 Level 1 DFD Admin

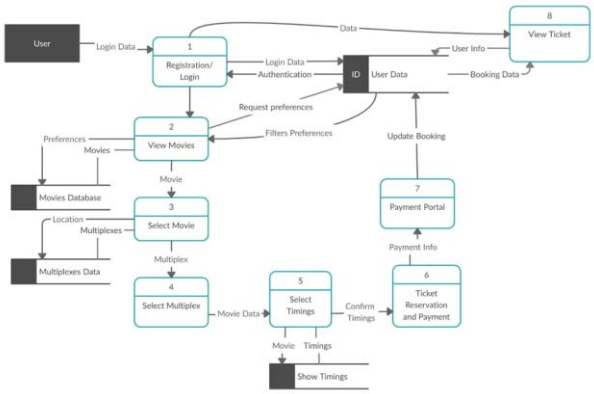


Fig.6 Level 1 DFD User

In the fig.5 and fig.6 the level 1 DFD is mentioned for the user and admin. In both the figures it is clearly noted the depth of the system for both the user and admin. It brings the efficiency to carry out the project

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:03:02 UTC from IEEE Xplore. Restrictions apply.

task. It makes the system understanding efficient and effective [14].

B. RECOMMENDER SYSTEM

The MMS will have a recommender system that trains a model on the user's previous ticket booking and cancellation history to predict the likelihood of the user purchasing tickets for a movie that is currently shown at the multiplex. Psychology of consumer behavior dictates that consumer are likely to avail/purchase products/services that are available to them right away. Therefore, it is intuitive that having movies that the user is more likely to book I the first screen he logs on to increases the chance of tickets being booked. This is especially true when users are quickly glancing at the homepage of the website prompting them to get tickets, whereas customers who use the service with a plan to book specific tickets to movies are not influenced by this as much [11]. We are proposing to use a logistic regression model as the base for the recommender and then build a more advanced model with greater accuracy. Logistic regression is a supervised classification algorithm. In a classification problem, the target variable (or output), y, can take only discrete values for a given set of features (or inputs), X. Contrary to popular belief, logistic regression is a regression model. The model builds a regression model to predict the probability that a given data entry belongs to the category numbered as "1". Just like Linear regression assumes that the data follows a linear function, Logistic regression models the data using the sigmoid function. Logistic regression becomes a classification technique only when a decision threshold is brought into the picture. The setting of the threshold value is a very important aspect of Logistic regression and is dependent on the classification problem itself. The features used in training the recommender are:

- Director
- Actors
- Genre
- Amount of ticket
- Language
- Subtitle-Language
- Time of the show
- Duration

C. APIs USED

- *Online Payment System*

As our system requires the facility for users to pay for the tickets, the book we plan to integrate a feature to pay online. To facilitate this, the system would require the use of already developed and tested Payment Gateways. Using Online Payment API like PayU, RazorPay, or the GooglePay API we can easily integrate it into our system. Configuring the API by using simple post requests the user will be transferred to the safe and secure payment gateway and on completion of the payment the user will be returned to our application.

As our system requires the facility for users to pay for the tickets, the book we plan to integrate a feature to pay online. To facilitate this, the system would require the use of already developed and tested Payment Gateways. Using Online Payment API like PayU, RazorPay, or the GooglePay API we can easily integrate it into our system. Configuring the APIs by using simple post requests the user will be transferred to the safe and secure payment gateway and on completion of the payment the user will be returned to our application.

1) Technologies used

The technologies that are being currently used to make MMS are:

- Desktop application – Java, C#, XAML
- Web application – Javascript, PhP, HTML, CSS
- Frameworks – Spring, Spring boot, React, Node.js
- Database – MySQL, MongoDB
- Diagram and Design tools - app.diagrams.net, Microsoft Project and Visio, NClass, Maven

There are various TMS and MMS available from different vendors available on different platforms and providing different sets of functionalities.

- Vista - Legacy software that has now started adapting to the latest technologies.
- CiniCloud - which provides everything from online ticketing to inventory and personnel management as well as cloud hosting (ticket commission).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:03:02 UTC from IEEE Xplore. Restrictions apply.

- Sony TMS – Most featured to manage multiplexes and internet Ticketing.

  Barco TMS – Mostly the screen/auditorium management.

The technology for the system can be based off of pre-existing technologies and aims to be more user-experience focused with the recommender system aiming to boost customer retention. The system will have many algorithms to suggest movie recommendations, to filter choices based on chosen preferences such as seating, location and other criteria.

V. RESULT ANALYSIS

The project implemented is a web-based user interface that has the options for both the user and admin. It is very much friendly to both of the user and admin. The frontend of the system is completely web based and interface for the user too. Recommender system is applied for the backend part with web partially used. Central databases are functionally used. Desktop applications like Java and C# with HTML, CSS and java script was used for the frontend option with node.js and spring boot framework.

VI. CONCLUSION AND FUTURE WORK

After completion of this project, Admins will be able to structure Screens/Auditoriums, add movies, schedule screenings, add or remove employees as well generate statistical reports [19]. Employees will be able to login and view, book, or cancel tickets. Users can sign up, sign in and view, book, or cancel tickets. The system can generate reports, manage to screen, and make personalized movie suggestions to users. And the actors will be able to edit some of their credentials according to permissions granted. The system cannot be linked to payment gateways and therefore all booking will be regarded as paid for and cancellations reimbursed immediately [17]. The system is limited to booking services for the users and does not offer any other services to them. The system cannot also integrate with movie distributors to project the movies on the screens. The system also does not handle tax-related tasks in accounting. Different newly tools and technologies are being developed day by day, the system can be made more efficient and can be structured again in good methods and design with high computation methods and techniques. We can also use deep learning technologies to enhance the model to the greater extend and bring the effectiveness into the system.

ACKNOWLEDGMENT *(Heading 5)*

We take this opportunity to express our profound gratitude and deep regards to firstly to Professor DR. Manoj Kumar for his exemplary guidance, monitoring, and encouragement throughout this project. His constant guidance and supervision helped a long way in this project's completion. Our thanks and appreciation also goes to our Department of Computer Science and Engineering, and Delhi Technological University for Helping Us to Develop this project.

REFERENCES

[1] O. F. O. R. E. Co-operation and I. E. Agency, "D Iffusion : T He C Ase of," *Energy*, vol. 87, no. October, pp. 645–659, 2006.

[2] Z. Wang *et al.*, "Failure prediction using machine learning and time series in optical network," *Opt. Express*, vol. 25, no. 16, p. 18553, 2017, doi: 10.1364/oe.25.018553.

[3] M. L. Popa, R. Albulescu, M. Neagu, M. E. Hinescu, and C. Tanase, "Multiplex assay for multiomics advances in personalized-precision medicine," *J. Immunoass. Immunochem.*, vol. 40, no. 1, pp. 3–25, 2019, doi: 10.1080/15321819.2018.1562940.

[4] P. Courlet *et al.*, "Development and validation of a multiplex UHPLC-MS/MS assay with stable isotopic internal standards for the monitoring of the plasma concentrations of the antiretroviral drugs bictegravir, cabotegravir, doravirine, and rilpivirine in people living with HIV," *J. Mass Spectrom.*, vol. 55, no. 6, pp. 1–13, 2020, doi: 10.1002/jms.4506.

[5] F. Shen, B. Sun, J. E. Kreutz, E. K. Davydov, W. Du, and R. F. Ismagilov, "Quantification of HIV and HCV viral load with large dynamic range using multivolume digital reverse transcription PCR on a rotational slipChip," *15th Int. Conf. Miniaturized Syst. Chem. Life Sci. 2011, MicroTAS 2011*, vol. 1, pp. 290–292, 2011.

[6] P. Szodoray, P. Alex, J. G. Brun, M. Centola, and R. Jonsson, "Circulating cytokines in primary Sjögren's syndrome determined by a multiplex cytokine array system," *Scand. J. Immunol.*, vol. 59, no. 6, pp. 592–599, 2004, doi: 10.1111/j.0300-9475.2004.01432.x.

[7] M. Cailleaux *et al.*, "Impact of a multiplex PCR assay (FilmArray®) on the management of patients with suspected central nervous system infections," *Eur. J. Clin. Microbiol. Infect. Dis.*, vol. 39, no. 2, pp. 293–297, 2020, doi: 10.1007/s10096-019-03724-7.

[8] J. P. Borde *et al.*, "Diagnosis of prosthetic joint infections using UMD-Universal Kit and the automated multiplex-PCR Unyvero i60 ITI® cartridge system: a pilot study," *Infection*, vol. 43, no. 5, pp. 551–560, 2015, doi: 10.1007/s15010-015-0796-4.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:03:02 UTC from IEEE Xplore. Restrictions apply.

[9] L. Bertazzi, G. Paletta, and M. G. Speranza, "Minimizing the total cost in an integrated vendor-managed inventory system," *J. Heuristics*, vol. 11, no. 5-6 SPEC. ISS., pp. 393–419, 2005, doi: 10.1007/s10732-005-0616-6.

[10] P. M. Valcour, "Managerial behavior in a multiplex role system," *Hum. Relations*, vol. 55, no. 10, pp. 1163–1188, 2002, doi: 10.1177/0018726702055010079.

[11] A. Alonso, *Chapter in DNA electrophoresis protocols for forensic genetics*, no. 2012. 2012.

[12] B. K. Shah, V. Kedia, and R. K. Jha, "Integrated Vendor - Managed Time Efficient Application to Production of Inventory Systems," pp. 275–280, 2021.

[13] B. K. Shah, A. Kumar, and A. Kumar, "Dog breed classifier for facial recognition using convolutional neural networks," *Proc. 3rd Int. Conf. Intell. Sustain. Syst. ICISS 2020*, pp. 508–513, 2020, doi: 10.1109/ICISS49785.2020.9315871.

[14] S. Adhikari, S. Thapa, and B. K. Shah, "Oversampling based Classifiers for Categorization of Radar Returns from the Ionosphere," *Proc. Int. Conf. Electron. Sustain. Commun. Syst. ICESC 2020*, no. Icesc, pp. 975–978, 2020, doi: 10.1109/ICESC48915.2020.9155833.

[15] M. H. Tania, K. T. Lwin, A. M. Shabut, K. J. Abu-Hassan, M. S. Kaiser, and M. A. Hossain, "Assay type detection using advanced machine learning algorithms," *2019 13th Int. Conf. Software, Knowledge, Inf. Manag. Appl. Ski. 2019*, 2019, doi: 10.1109/SKIMA47702.2019.8982449.

[16] H. Ye, L. Liang, G. Y. Li, J. B. Kim, L. Lu, and M. Wu, "Machine learning for vehicular networks," *arXiv*, no. April 2018, 2017.

[17] C. From, "O Vercoming I Nterference in S Patial," *Ieee Wirel. Commun.*, no. December, pp. 95–104, 2007.

[18] D. Ma, W. H. Ki, C. Y. Tsui, and P. K. T. Mok, "Single-inductor multiple-output switching converters with time-multiplexing control in discontinuous conduction mode," *IEEE J. Solid-State Circuits*, vol. 38, no. 1, pp. 89–100, 2003, doi: 10.1109/JSSC.2002.806279.

[19] L. Buitinck *et al.*, "API design for machine learning software: experiences from the scikit-learn project," pp. 1–15, 2013, [Online]. Available: http://arxiv.org/abs/1309.0238.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 06:03:02 UTC from IEEE Xplore. Restrictions apply.

IOP Conference Series: Earth and Environmental Science

PAPER • **OPEN ACCESS**

Natural Microalgae Biofilm As A Tool To Clean-Up Mining Wastewater

To cite this article: S. Anbu Kumar *et al* 2021 *IOP Conf. Ser.: Earth Environ. Sci.* **775** 012014

View the article online for updates and enhancements.

This content was downloaded from IP address 14.139.251.100 on 07/07/2021 at 07:34

Natural Microalgae Biofilm As A Tool To Clean-Up Mining Wastewater

S. Anbu Kumar[1] Shivam Gupta[2] Shubham Kumar[3] Pawan Kumar Meena[4]

[1] Civil Engineering Department, Delhi Technological University, Delhi, India
[2] Civil Engineering Department, Delhi Technological University, Delhi, India
[3] Civil Engineering Department, Delhi Technological University, Delhi, India
[4] Civil Engineering Department, Delhi Technological University, Delhi, India

E-mail : [1] sanbukumar@dce.ac.in [2] shivamdtu0908@gmail.com [3] shubham2798@gmail.com [4] pkmcricket02@gmail.com

Abstract. Pollutants in amphibian conditions will in general change the network arrangement and exercises of biofilm cells. The making of excessive reactive oxygen species (ROS) which is caused due to the stress induced out of these pollutants play a major role in debasement of pollutants in oceanic conditions. This smaller than expected audit is pointed toward anticipating the proficiency of the utilisation of regular biofilm in water treatment coming about because of mining exercises. Additionally, it sums up the idea of mining by briefly expressing the stages engaged with mining and how the toxins delivered during mining can make harmful impacts on amphibian vegetation. Discussion on how the advancement of biodiversity is impacted by the ecological issues resulting from activities of mining was done. Microalgae assume a significant part in the alleviation of the effect of contamination coming about because of mining exercises. The cycle of toxins expulsion in freshwater by miniature green growth biofilm is by debasement and biosorption. Biosorption of hefty metals in mining wastewater was caused because of the extracellular polymeric substances (EPS). The kinetic model should be understood properly which explains the biosorption of hefty metals through physical and chemical reactions.

1. INTRODUCTION

Mining is the principal monetary action of many non-industrial nations. Mining was among the significant wellspring of fare in Nigeria before 1970. The over reliance on oil assets has strengthened distinctive mining practice in the nation because of absence of government will, mediation and usage of accessible guidelines. Mining in Nigeria is managed without thinking about the effects on the climate, and as a rule, recovery is never a phase in mining measures. Mining activity, be it little or huge scope, is intrinsically troublesome to the climate, delivering tremendous amounts of waste that can have a malicious effect for quite a long time. The danger of contamination from mining destinations keeps going even when the mining exercises are completed. Natural issues exist where mining exercises are done.

These issues show as corruption of woodland and biodiversity loss, tainting of soil, contamination of air and groundwater, clamor and vibrations, decay of common waste frameworks, regular land debasement, negative effect on earthbound and amphibian biological systems, financial and human


Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

wellbeing. The deteriorating climate brought about by mining happens mostly because of inadmissible and inefficient working practices and restoration measures. Subsequently, ensuring water quality is a pressing ecological test at mining and mineral preparing locales [1]-[3].

Explanation, vanishing, crystallisation, sterilisation, particle trade, film division, bundle plants and organic methods are some of the numerous advances utilised in wastewater treatment. These innovations are cost-insufficient and hard to use under field conditions, consequently in such a circumstance, there is an earnest need to consider common, straightforward and savvy procedures for control. The utilisation of common microalgae biofilm for cleaning of water coming about because of mining exercises is an option in contrast to ordinary and other natural strategies because of its cost-adequacy, the necessity of low innovative skill and good proficiency to eliminate supplements, weighty metals, natural and inorganic wastewater from wastewater.

Inorganic surfaces in relationship with normal freshwater microalgae biofilm envelops a few syntrophic gatherings of microorganisms. Foul extracellular framework encircles part cells shaping a biofilm which is made out of extracellular polymeric substances (EPS). EPS is delivered through aggregation of lipid, proteins, DNA of every individual framing characteristic microalgae biofilm. Network routine of part microorganism is implied by the parts of EPS in microalgae biofilm and possess 3D structure [4].

The microorganisms creating freshwater microalgae biofilm are physiologically not the same as other planktonic cells of a comparative living being coasting in a freshwater climate. Common biofilm comprises various sorts of miniature organisms, these either take care of autotrophically or heterotrophically. Autotrophic life forms present in common biofilms comprise miniature green growth and cyanobacteria, though heterotrophic living beings incorporate microorganisms, parasites and protists. A characteristic biofilm is structured by these organisms encircled by EPS on outside of either a fake or normal substrate. The EPS of characteristic biofilm is the motivation behind why biofilm can oppose natural pressure, consequently causing the debasement of poisons in sea-going conditions. This smaller than normal audit is subsequently pointed toward estimating the productivity of the utilisation of characteristic microalgae biofilm for treatment of water coming about because of mining exercises [5].

2. THE CONCEPT OF MINING

Mining can be characterised as the cycle by which minerals coming about because of geographical cycles are removed from the lithosphere. These monetarily important minerals are a mix of various materials which include coal, metals, chalk, measurement stone, potash, dirt etc. The way toward mining includes the extraction of non-sustainable assets, for example, coal, unrefined petroleum and flammable gas which is autonomous of material originating from agronomic exercises and also the materials made in research facility. The unanswered inquiry regarding mining activities is its negative natural effect which has prompted the production of various laws pointed toward diminishing the negative part of mining on the climate. The stages associated with mining exercises incorporate prospecting, exploration, development, exploitation and reclamation as shown below in Figure number 1 [6].

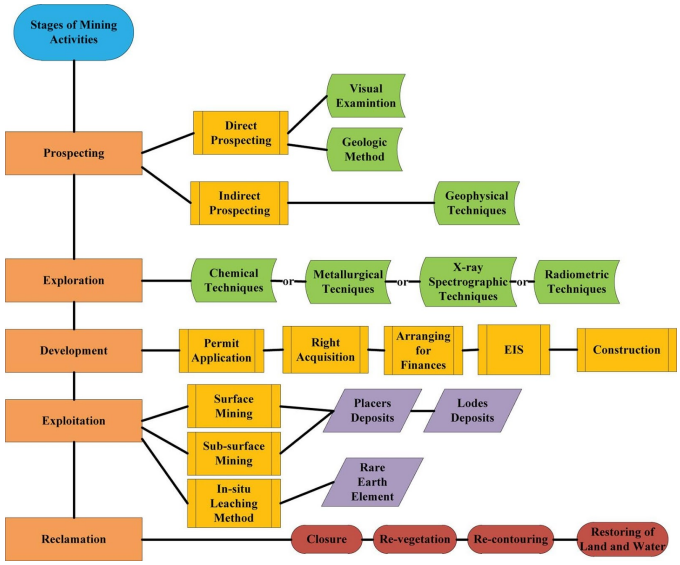


Figure 1. The phases occurring in mining process

3. ENVIRONMENTAL IMPACTS DUE TO MINING ACTIVITIES

Natural issues coming about because of mining exercises are because of the immediate or circuitous effect of mining rehearsals. These issues are of nearby, provincial and worldwide concern and emerged from the emanation during mining exercises. These outflows prompted ecological issues, for example, sinkholes, disintegration, soil pollution, water tainting (surface and groundwater) (as shown in Figure number 2), extinction of species and weakening of quality of air. The way toward mining is related with soil debasement and disintegration prompting the obliteration of enormous territories of land. Disintegration and debasement are because of mining exercises, for example, open pit and overburden mining, development, and so on Soil corruption is because of the tainting of soil by the substance toxins created during mining exercises. One significant impact of soil disintegration coming about because of mining exercises is relation with water assets defilement and decrease in variety of life. The contaminations created during mining exercises can saturate springs and dirty subsurface water or dissolve within close by run-off water prompting increment in contamination, stream sedimentation, obstructing of streams and loss of biodiversity [7]-[8].

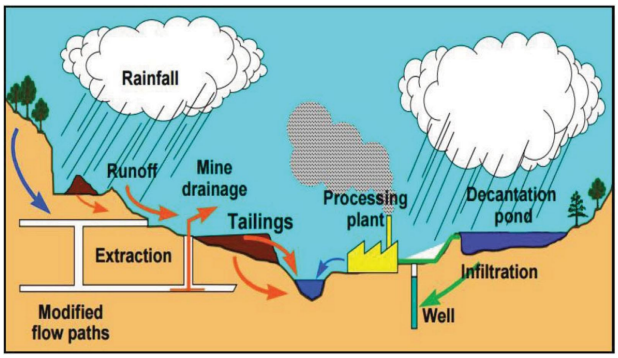


Figure 2. Effect on surface and groundwater due to mining activities

4. THE EFFECT ON BIODIVERSITY DUE TO MINING ACTIVITIES

Change of atmosphere coming about because of mining exercises is related with unsettling influence and determination antagonistic wellbeing impacts on both amphibian and earthly creatures. Harm to the first site is more negative to biodiversity in light of the fact that the toxins delivered during mining exercises are poisonous and can straightforwardly or in a roundabout way influence biodiversity.

The alteration of the territory is related according to variation in pH and also the climate which thus makes unsettling influences the encompassing network. Cosmopolitan species are probably going to adjust according to the adjustment in pH and climate. On the other hand, indigenous species are influenced because of slight variety in the state of their environment which builds the danger of annihilation [9]-[10].

5. THE EFFECT ON AQUATIC ORGANISMS DUE TO MINING ACTIVITIES

Mining affects the amphibian climate in a total unexpected way. One of these includes immediate harming; which happens when toxins are bioavailable in the water or potentially openly portable in the silt. Toxins can genuinely influence sea-going creatures. Models remembering high groupings of suspended residue for a stream will in general restrict entrance of light due to which algal biomass is lessened. Biomass is limited because of metal oxide oath through covering green development or substrate of theirs, in like manner hindering colonisation. Likewise, temperature, precipitation, pH, saltiness, metal fixations in a flow could over the lengthy haul influence oceanic networks. Change in pH or climate influence dissolubility of metal which makes it even more bioavailable so that living beings can take it. It is important that metal impurities persevere in the environment for quite a while as long as 90 years after pyrite mine conclusion. Mine wastes cause trouble in algal networks prompting diminished essential efficiency. Both physical and compound changes in sea-going conditions will in general influence diatoms populaces [11]-[12].

6. POLLUTION IN FRESHWATER SURROUNDINGS DUE TO MINING ACTIVITIES

The poisons delivered while performing mining exercises happen to be natural and inorganic. Natural contaminations are isolated in the direction of two divisions against extremity namely hydrophilic and hydrophobic natural poisons. The polar natural impurity which can be dissolvable with water can be considered hydrophilic contamination, on the other hand non-polar including insoluble natural pollutant is known as hydrophobic natural foreign substance. Hydrophilic impurities are natural toxins that don't promptly influence freshwater environment attributable toward this non-industriousness value and capacity of breaking up into the water, regular illustration to hydrophilic foreign substance incorporates natural contaminations with moderately little alkyl gathering, for example, liquor and carboxylic acids (CH3, C2H5, C3H7 and C4H5). Another significant model is named methyl tertiary butyl ether (MTBE). Hydrophobic toxins are the most troubling natural impurities because of their diligence's tendency and failure to break down in water, illustration of these poisons incorporates dichloro-diphenyl-trichloroethane (DDT), polychlorinated bi-phenyls (PCBs), poly-brominated di-phenyl ethers (PBDEs), polychlorinated di-benzodioxins (PCDDs), polychlorinated di-benzofurans (PCDFs), poly-sweet-smelling hydro-carbons (PAHs), benzene toluene, ethylbenzene and xylene (BTEX). Inorganic toxins incorporate substantial metals, phosphorus mixes, nitrogen mixes, halide mixes, sulphide mixes, and so on [13]-[14].

Hefty metals can be considered as utmost tenacious inorganic impurities because they have non-biodegradable limit. By and by, these components may be found in a low focus in the climate yet they

will in general be hindering amphibian greenery because of their capacity to bioaccumulate along the evolved way of life. The compounds which have phosphorous and nitrogen are likewise troubling to climate because of their commitment to eutrophication that prompts developments of unsafe green growth. Destructive microalgae can be hazardous to new water living space because of the development of green growth sprouts. Microalgae blossoms change in shading from green, yellow or earthy coloured tone contingent upon the types of miniature green growth included. This sprout is caused because of unnecessary use of supplements by hurtful miniature green growth especially in sluggish water bodies, for example, lakes and lakes.

These sprouts created through freshwater microalgae are incredibly poisonous for freshwater biota since sprouts will in general hinder the fishes' gills and further freshwater macro-invertebrate also exhausts the accessible broke down the oxygen present into the freshwater natural surroundings prompting the demise of beings living in freshwater. Blossoms can likewise make a biological system to be impractical in light of the fact that they will in general be toxic to the more modest fishes plus all the channel feeder living beings which feeds on aquatic living beings. This poisonousness penetrates the food web coming about through passing of full scale plants in natural water conditions. [15]-[16].

Destructive microalgae can likewise influence sea-going profitability since they keep daylight energy from infiltrating profound through alternative freshwater belts prompting decrease of variety of living beings in freshwater territory.

The counteraction of daylight by unsafe miniature green growth can be founded on staining of freshwater territory through destructive microalgae or because of arrangement of broad layer on outside of freshwater by freshwater microalgae. Unsafe microalgae blossom likewise influences the versatility of freshwater fauna, for example, coral and can likewise influence other freshwater greenery from approaching daylight for photosynthesis since they lowered this freshwater verdure and furthermore fill in as a shade, consequently keeping daylight energy from arriving at the life forms. Hurtful microalgae can create mycotoxins that affect freshwater ecology especially fishes along with other miniature plus large scale spineless creatures.

7. ESTABLISHMENT AND ARRANGEMENT OF NATURAL FRESHWATER MICROALGAE BIOFILM

Freshwater microalgae are those microbes which are adjusted with a freshwater climate. These microalgae are known to be unicellular and free living beings whose size ranges from about 4-178 micrometer. Microalgae can use CO2, water and light to perform photosynthesis so that they get the energy needed for metabolic exercises. Development of microalgae likewise relies upon significant supplements, for example, nitrogen and phosphorus. Normal freshwater microalgae biofilms are shaped from connection between free-skimming microalgae to natural substrate. The underlying new water microalgae which structures biofilm can connect with the substrate from outset with assistance of powerless Van der Walls power of fascination and hydrophilic impact prior to join forever to expanding thickness of the freshwater microalgae. The cycle of biofilm development with the aid of freshwater microalgae conceivably isolated in five phases. These phases incorporate starting connection, lasting connection, beginning development, perpetual development and scattering [17].

The phases of regular freshwater biofilm spoke to in Figure 3 portray that during the underlying and perpetual connection arranges, the EPS was not shaped because of the perceivability of the substrate, however the quantity on the substrate of freshwater microalgae during the underlying connection phase is less as compared to the lasting connection stage. The EPS was shaped during the underlying development, lasting development and scattering stage, regardless of the broadness of the characteristic freshwater biofilm of perpetual development and scattering phases is broader when

compared to underlying development phase. The reason behind greater biomass of the freshwater microalgae were found during the development and scattering stages when contrasted with the connection stage is inferable from the way that when in connection phase, small amount of freshwater microalgae couldn't secure with substrate rather it appended to the EPS during the development and scattering stages. The scattering phase is last phase of freshwater biofilm microalgae arrangement. It is the phase where freshwater microalgae biofilm sets up and might go through structural and size adjustments. Proteins, for example, dispersin B assume a significant part for diffusion of freshwater microalgae biofilm in light of the fact that they can debase the EPS. These freshwater biofilms get typically discovered normally upon strong substrates lowered with amphibian conditions, despite the fact that they can here and there shape a drifting mat on the surfaces of sea-going living space. Microalgae biofilm regularly develops to a plainly visible level however will in general contain different sorts of microorganisms incorporating microbes, protozoa, archaea and growths alongside every gathering of living being assuming some specific metabolic job [18].

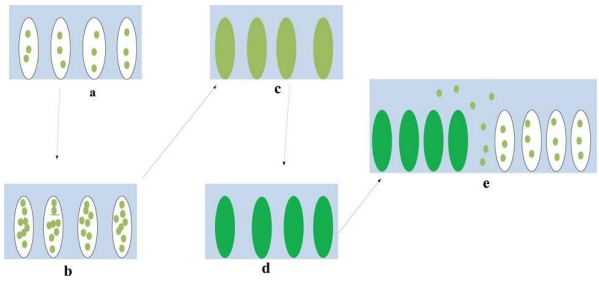


Figure 3. Biofilm formation process :

primary attachment
lasting attachment
primary maturation
lasting maturation
dispersion.

8. NUTRIENT AVAILABILITY

For development of biofilm normally, new water microalgae used the supplements that are available in amphibian climate, for example, nitrate. Despite the fact that, during the early development phase of freshwater microalgae, the phones are metabolically dynamic and will in general use a greater number of supplements as compared to late development phase where it is metabolically little dynamic.
This guideline will in general be inconsequential for normal freshwater biofilm development in light of the fact that during the development phase, the biological film isn't yet stabilised dependent on less thickness and also nonattendance of EPS, along these lines spending lesser measure of supplements contrasted with later development phase where the microalgae becomes completely developed, which settles serenely to the natural substrate because it has higher thickness. The arrangement of EPS additionally adds to the high convenience of supplements on the grounds that the EPS helps the connection of other non-microalgae life forms which likewise need supplements to endure. Studies have indicated that raised supplements focus in amphibian conditions will in general advance EPS creation while supplement starvation prompts the corruption of EPS bringing about the unit of the contributing creature in the biofilm arrangement. The reason behind denial of biomass of microalgae to utilise in creation of lipid is due to huge measure of supplements spent by the microalgae biofilm in

later phase during development. On the other hand, it also braces the inspiration that sea-going bodies consisting of microalgae biofilm don't allow green growth sprouts.

9. EXTRACELLULAR MATRIX (EPS) OF MICROALGAE BIOFILM

To give biochemical and underlying scaffoldings which build the freshwater biofilm, EPS grid is used. It is a 3-D organisation comprising various acids. The section of this EPS in biofilm affiliation is used for encasing the phones, also furthermore help during interchanges, biological signs and quality trade. The EPS additionally assumes a function in the outer processing of toxins on the grounds that the lattice traps extracellular compounds and keeps them in closeness to the member cells.

Freshwater microalgae with biofilm affiliation will probably have various attributes for free-gliding microalgae because there is some connection happening mutually with the microalgae and creatures inside film's thick and ensured climatic conditions. Motivation behind why the microalgae is profoundly impervious when we talk about cleansers or antitoxins is on grounds that the thick extracellular framework and the external layer of cells secure the inside of the network [19].

10. FUNCTION OF NATURAL FRESHWATER MICROALGAE BIOFILM IN DEGRADATION OF MINING WASTEWATER POLLUTANTS

Mining exercises result in creation of wastewater. This affects the freshwater flora. Due to characteristic rule during the creation of this microalgae biofilm, some dangers exist and then this new water microalgae film is used to treat mining wastewater. This assists in limiting that danger, and it makes it effectively open and financially savvy.

The instruments associated with poison corruption by microalgae biofilm are a characteristic cycle including biosorption and connection. Since a large number of microalgae cells are associated with microalgae biofilm development, these biofilms can biodegrade or biotransform contaminations coming about because of mining exercises by one or the other debasement or amassing. The capacity of microalgae biofilm to bring about the corruption in poisons that exist with mining wastewater can be because of the protein exercises created by the individual microalgae cells in the biofilm or because of the expanded degree of responsive oxygen species (ROS) delivered because of the high measure of microalgae cells shaping the biofilm.

Numerous examinations have indicated that freshwater microalgae biofilm can create responsive oxygen species (ROS) which is equipped for oxidising steady natural compound, consequently achieving the debasement of these mixes into CO2, CO32− and H2O with the items being used by the freshwater microalgae for biomass creation. Studies show that freshwater microalgae biofilm can create ROS of high focus, which is equipped for corrupting determined natural mixes, for example, tetra-chlorobenzene, alkyl halides, trichlorobenzene or some substantial natural mixes, for example, PCBs. These examinations additionally uncover a fact which says the ROS created by this microalgae could be equipped for debasing toxins that contain sulphur in light of the possibility of ROS to desulfurize thio-carbonyl mixes, for example, thiouracil, thioamides and thioureas to carbonyl mixes. Freshwater microalgae have cell association like microbes yet are distinctive about their capacity in carrying out the process of photosynthesis. This ability of freshwater microalgae that can carry out photosynthesis can be critical for existence on land and water on the grounds that during the photosynthetic cycle, they are destined to debilitate the ozone harming substances especially CO2 created during mining exercises prompting the creation and arrival of more oxygen.

Microalgae biofilm can eliminate inorganic sub-positions, for example, phosphorus, alkali and nitrogen exacerbates existing within freshwater natural surroundings through biomass creation with

the help of these materials, subsequently keeping the material from amassing to some harmful stage within living space of freshwater. Freshwater microalgae are likewise ready to oxidise or decrease poisonous metals in freshwater conditions, accordingly forestalling the poisonousness to freshwater biota. This is likewise conceivable through the creation and arrival of ROS by freshwater microalgae which oxidises these harmful metals and turn them into operable components. Investigations of marine climate demonstrate that ROS created through marine microalgae would be equipped for expanding the measure of iron which would exist in climate through a little decrease of ferric mixes [20].

11. CONCLUSION AND FUTURE DIRECTION

Mining exercises adversely change oceanic as well as earthly environments hugely. The water created will in general influence the flora and fauna of getting surface water and furthermore represent some basic danger for general wellbeing of earthbound living beings. The perpetual high quality mining exercises in non-industrial nations are troubling because of the danger related with it. Most high quality mining is completed without natural effect appraisal (EIA), and accordingly, the destiny of the ecological contaminations created during mining exercises isn't thought of. These informal mining exercises are the reasoning behind lead harming in three neighbourhood government zones in Nigeria in March 2010.

Microalgae biofilm is also utilised like an instrument in refurbishing of water contamination coming about because of mining exercises. This is on the grounds that microalgae biofilm can happen normally in sea-going conditions, and hence its utilisation with respect to field conditions can never be a troublesome undertaking. Improvement shown by natural factors, for example, light and supplements should help in development of microalgae biofilm. Microalgae biofilm will in general forestall the events of green growth sprouts which is the serious danger to the freshwater climate. The reproductions of natural elements can likewise motivate this microalgae biofilm to deliver a greater measure of receptive oxygen species (ROS) which may control the development of microorganisms, in this manner limiting the deficiency of broken down oxygen.

Despite the fact that, microalgae biofilm eliminates poisons present in water resulting from mining wastewater by corruption and biosorption. Still, further exploration is needed for explaining whether the proteins delivered through microalgae biofilm in the course of the corruption cycle are the sole commitment of the whole cells participating in biofilm or that of the extracellular lattice. Additionally, part of EPS in cycle of biosorption of substantial metal expulsion from water resulting from mining needs to be investigated and also get clarified about its evacuation instrument. It is likewise imperative to extensively comprehend the motor model which clarifies the synthetic and actual responses of hefty metal assimilation from outside of this microalgae biofilm.

References

1. Xie F, Zhang TA, Dreisinger D, Doyle F, "A critical review on solvent extraction of rare earths from aqueous solutions", Miner Eng. 56:10–28, 2014.
2. H.Palma, E.Killoran, M.Sheehan, F.Berner, K.Heimann, "Assessment of microalga biofilms for simultaneous remediation and biofuel generation in mine tailings water", Bioresource Technology, vol. 234, pp. 327-335, 2017.
3. Neha Mishra, Neetu Mishra, "Utilization of Microalgae for Integrated Biomass Production and Phycoremediation of Wastewater", J. Algal Biomass Utln., 8(4): 95-105, 2017.
4. Marziyeh Molazadeh, Hossein Ahmadzadeh, Hamid R. Pourianfar, Stephen Lyon, Pabulo Henrique Rampelotto, "The Use of Microalgae for Coupling Wastewater Treatment With

CO2 Biofixation", Front. Bioeng. Biotechnol., 2019.
5. Bessemer AL, "Diversity, community structure and function of biofilms in stream ecosystems", Res Microbiol. 166(10):774–781, 2015.
6. Felix Wollmann, Stefan Dietze, Jörg-Uwe Ackermann, Thomas Bley, Thomas Walther, Juliane Steingroewer, Felix Krujatz, "Microalgae wastewater treatment : Biological and technological approaches", Eng Life Sci. 19:860–871, 2019.
7. Mohamad Shurair, Fares Almomani, Simon Judd, Rahul Bhosale, Anand Kumar, Ujjal Gosh, Majeda Khreisheh, "Advanced wastewater treatment using microalgae: effect of temperature on removal of nutrients and organic carbon", 7th International Conference on Environment and Industrial Innovation, 2017.
8. Hammouda, O., Gaber, A., Abdelraouf, N, "Microalgae and Wastewater Treatment", Ecotoxicology and Environmental Safety, vol. 31, Issue 3, pp. 205-210, 1995.
9. Fica, Z.T., Sims, R.C., "Algae-based biofilm productivity utilizing dairy wastewater: effects of temperature and organic carbon concentration", J Biol Eng 10, 18, 2016.
10. N.Abdel-Raouf, A.A.Al-Homaidan, I.B.M.Ibraheem, "Microalgae and wastewater treatment", Saudi Journal of Biological Sciences, vol. 19, Issue 3, pp. 257-275, 2012.
11. Ana F. Miranda, Narasimhan Ramkumar, Constandino Andriotis, Thorben Höltkemeier, Aneela Yasmin, Simone Rochfort, Donald Wlodkowic, Paul Morrison, Felicity Roddick, German Spangenberg, Banwari Lal, Sanjukta Subudhi, Aidyn Mouradov, "Applications of microalgal biofilms for wastewater treatment and bioenergy production", Biotechnol Biofuels, 10: 120, 2017.
12. Ansari, A.A., Khoja, A.H., Nawar, A., "Wastewater treatment by local microalgae strains for CO2 sequestration and biofuel production", Appl Water Sci 7, pp. 4151–4158, 2017.
13. Navarro, A., Collado, D., Carbonell, M., "Impact of mining activities on soils in a semi-arid environment: Sierra Almagrera district, SE Spain", Environ Geochem Health 26, pp. 383–393, 2004.
14. Satpal, Satpal, "Waste Water Treatment Using Microalgae -A review Paper", 2016.
15. Aboka Yaw Emmanuel, Cobbina Samuel Jerry, Doke Adzo Dzigbodi, "Review of Environmental and Health Impacts of Mining in Ghana", J Health Pollut., 8(17): 43–52, 2018.
16. KAUR, P., BANSAL, A., NARAYAN, O., "MICROALGAE FOR RECTIFYING WASTE WATER: A REVIEW", Journal of Biology and Nature, vol. 11, Issue 3, pp. 9-14, 2020.
17. Gulshan Kumar Sharma, Shakeel Ahmad Khan, "Bioremediation of Sewage Wastewater Using Selective Algae for Manure Production", International Journal of Environmental Engineering and Management, vol. 4, no. 6, pp. 573-580, 2013.
18. Karel Prach, Anne Tolvanen, "How can we restore biodiversity and ecosystem services in mining and industrial sites?", Environmental Science and Pollution Research, vol. 23, Issue 14, pp. 13587-13590, 2016.
19. M.A. Riaz, B. Ijaz, A. Riaz, M. Amjad, "Improvement of waste water quality by application of mixed algal inocula", Bangladesh J. Sci. Ind. Res., vol. 53, Issue 1, pp. 77-82, 2018.
20. Dalrymple, O.K., Halfhide, T., Udom, I., "Wastewater use in algae production for generation of renewable resources: a review and preliminary results", Aquat. Biosyst. 9, 2, 2013.

2021 2nd International Conference for Emerging Technology (INCET)
Belgaum, India. May 21-23, 2021

Object Tracking in a Zone using DeepSORT, YOLOv4 and TensorFlow

Dr. Rajni Jindal
Delhi Technological University,
Delhi, India
rajnijindal@dce.ac.in

Aditya Panwar
Delhi Technological University,
Delhi, India
adityapanwar_2k17co27@dtu.ac.in

Nishant Sharma
Delhi Technological University,
Delhi, India
nishantsharma_2k17co214@dtu.ac.in

Aman Rai
Delhi Technological University,
Delhi, India
amanrai_2k17co45@dtu.ac.in

2021 2nd International Conference for Emerging Technology (INCET) | 978-1-7281-7029-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/INCET51464.2021.9456443

Abstract*: In this paper, we are watching out for the issue of different article track-ing in a single packaging. To oblige the course of action, we at first recognize all of the articles in the packaging and subsequently designate wonderful ID's to all of them until they move out of the edge of reference. We proposed to deal with this issue in three stages, which are according to the accompanying: perceiving all of the articles in the current edge followed by identi-fying them, and thereafter finally, we intend to follow them. We took this issue announcement considering the way that in spite of the way that we have various estimations identifying with this situation, the MOT-A figure is not adequate. So to manage on something practically the same, we hope to use Kalman channels for redesigned results and differentiated our MOT-An estimation and right now open works in this space and the results were adequate. The key computation used by us was YOLO after which the articles are recognized and organized under various classes. After this, for settled mo-tion figure and feature extraction, Kalman channels were used to show these states and as of late perceived articles are settled of a current edge, behind the linking which is created for new acknowledgment. It is done using the Deep SORT algo-rithm, which is mainly a Deep alliance metric using a SORT computation. We used kalman channels because they improve the precision of our model and besides redesigned the result for a most part. On practically identical lines, we have used YOLO to do entity disclosure as well as ID, both at the similar timing. It can in like manner be treated as an identifier, which when is applied a lone neu-ral association can expect the hopping box and do multiple class gathering. Our proposed plan can have distinctive multi-reason uses in our consistently life like it can hold people back from getting kept at a singular spot, especially during COVID times and rais-ing the alert to caution people. We can similarly use this approach to manage address prob-lems concerning manage the chiefs and perception zoo/biodiversity stops if there ought to emerge an event of contentions between different animals where checking each and every transgression gle area is no basic task.

Keywords: Kalman Filters, Multiple Object Tracking, SORT, YOLOv4.

I. INTRODUCTION

The issue of Multiple article following is a standard issue since today is the need vital. This is because consistently, we see various articles everywhere, be it roads, parks, film passages or markets. We are endeavoring to handle this issue using the top tier development with least information sources and endeavor to achieve the best results. Regardless, saying this and doing these things are two particular things. It is fairly perplexed, and it may make a greater number of counterfeit positives. For the abatement of problematic acknowledgments, we have earlier needed to use simply the SORT computation, but at this point we will overall use the Deep SORT estimation which uses Kalman channels for better gauge [1]. We have surveyed our model on the MOT18 dataset, with the progress in this field of learning estimation and colleague of new pos-sibilities with our regular multi-object following issue and besides to think about a couple of changes in the development front. What about we start with the beginning of such models. As of now we will analyze in a word about the MOT issue and we ought to appreciate what absolutely this MOT issue is? Different Object Tracking (MOT) issue can be simply put as looking at the picture once and perceiving all are the articles in the pic-ture. We go over various articles in a singular edge and perceiving is the first and the chief development, which requires a huge load of challenge and a while later after is another task. So we devised disconnecting the issue into two segments: acknowledgment and connection. The model is furthermore prepared to perform perception and moving article acknowledgment and following consistently.

By and by let us bounce into how is object following and revelation being performed? In this computation, we have used the SORT algo-rithm with a development, for instance the deep association grid is used that improve the organizing with procedure and reduce the amount of character switche by using the procedure which is to add visually appearing descriptor as well as appearance features. This part includes 2 segments which are first thing recognizing the things and a while later causing a leaping to limit round it every bit of packaging of the system and thereafter following the articles in the ricocheting box till it moves out from the degree of the packaging. Regardless, it was one critical issue which the philosophy due to which results were not adequate and the clarification was hindrance, i.e., our structure couldn't perform well when there was an article in the perspective on another thing since SORT relies upon a fundamental development code and it couldn't work outstandingly as indicated by needed results. Subsequently, we picked the Deep SORT computation for the same. Hence, it had the choice to per-structure better and give needed results. We could moreover betters the precision by using development assessment and through appear incorporates too.

For following of vehicles, we used the segregation estimation. It addresses Simple Online Real-time i.e. SORT which follows the going with progresses:

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:50:46 UTC from IEEE Xplore. Restrictions apply.

- It makes use of Hungarian strategy and kalman filtering to manage development predict as well as data alliance
- Entity model contains scale, entity position, skipping boxes extent as well as development gauge for the accompanying edge
- Arranging tends to the data relationship by finding out bobbing Intersection box over affiliation (IOU) distance
- The last errand is done based on the loU distance
- It has high precision and accuracy and is speedy and fundamental

How might we show up toward the completion of perceiving any article? Starting, an all out game plan of things in the packaging of the picture are recognized. Picture division is moreover an im-portant part of thing detection[2]. At the point when the disclosure is more than, an organizing with plan is set performed is done for tantamount acknowledgments inside the edge of the skipping box and they are followed till they stay inside it. The accompanying request arises concerning how are we getting along object following dynamically? For that, we from the outset made a lot of article detec-tions. By then we made a wonderful ID for all of the acknowledgments. After this, we followed the article all through some time frame till it stays in the packaging. Finally, we kept up the ID till the thing is in the packaging. At the point when it moves out of the edge, the article's ID was vanished, and this is how we are following the things.

Now let us understand the work of the Deep SORT technique as well as YOLO:

- Object ID as well as affirmation happen through YOLO Technique, for instance bouncing boxes are first made and a while later they are put under the recognized classes.
- Motions Predicting and features age happen which an appraisal model is made and kalman channels are use to exhibit this state's so we can follow them till their augmentation in the edge.
- Finally, following occurs with the help of Kalman stations in the past diagram and as of late recognized things are set in the current packaging, after which a connection is made for new revelation and this is cultivated for discovering them.

We utilized Kalman channels as they improved the exactness of our model as they yielded better outcomes. Our calculation can besides go presumably as a locater, which by apply a particular neural affiliation can expect the ricocheting box and can perform multiple class gathering. On equivalent line, that utilized YOLO can perform object ID and confirmation meanwhile. Further, we will talk about past related work in the field prior to examining our procedure and clarifying the idiosyncrasy in the equivalent.

II. Related Work

Past works has shown that the Neural Networking is proving to have a surprising computation for faster acknowledgment and request of authentic articles. Some typical applications of CNN are Decoding Facial Recognition, recognizing vehicles, Ancient and Environment groups, and vehicles as well as fauna on roads [3]. CNN modelling is called Single Shot Multi-box. Detecting make the recognizable proof communication essentially speedier [4].

For making a precise neural association, the going with progress is following:

- Capturing of picture
- Pre-setting up the image
- Segmentation of picture
- Extraction of features
- Distinguishing the image - accomplice the image which the class to has a spot

YOLO model as of late has gotten mainstream for object location and it is not quite the same as different strategies as it utilizes a solitary CNN that can predict a diverse bouncing box around the article into a picture and discovers different probabilities for partner various classes to various jumping boxes. Profound Learning is one of the calculations that has demonstrated its significance over and over in the space of item recognition, however YOLO improve induction rate for the identification of genuine articles as well as it makes the article discovery favorable to cess quicker with colossal exactness and continuous vigor and more reasonable to use, in actuality, scenarios[5].

In the MOT which is Multiple Object Tracking issue we endeavor to screen multi-ple objects that are there in the different edges in an image course of action. Pushing ahead to the second supporter of the problem, i.e entity following which in it-self is a troublesome endeavor, in fact, similarly as during video assessment where we need to screen a thing in different resulting edges by making a bobbing box around it and allocate an intriguing character to the article across all of the edges of that comparable thing which appears in the image gathering which can be either consistent or can be recorded video. It is easier when we have quite recently each article in the edge thusly anyway the issue becomes flighty when we have various things at a comparable second in an edge. Saying perceives all of the things in the packaging and apportions an exceptional character to each dissent and uses these characters to screen different articles in the packaging over time[6]. The article track-ing system has a stream wherein they play out the task of discovering various things of an image course of action examination which can be either steady or recorded video. First thing an article is perceived and a while later the improvement of the thing is continued in following frames[7].

Here in this endeavor, we do the thing tracking machine to distinguish multi-ple entities in an area inside an edge of picture game plans and keep the incorporate of the outright article in the edge similarly as the hard and fast item in a particular zone demonstrated [8]. Thusly, its show can be smoothed out beginning to end clearly and hence handling the info data quickly. For Tracking, we use the Extension of SORT methodology which is Simple Online and Realtime Tracking and when it is getting together with Deep Association Metric called Deep SORT estimation. To recognize and follow various things dependent on various qualities of the articles like the size of article, the speed of thing, shape and genuine appearance of the article, etc we

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:50:46 UTC from IEEE Xplore. Restrictions apply.

use YOLO and Deep SORT[9]. Kaman segregators are used when there is a hindrance in the edge. Kalman filters work on consistent speed and straight discernment and assessments the method of the article when that thing isn't perceptible due to blockages like smoke, murkiness, flying bird, etc In entity disclosure as well as worldwide situating systems, Deep SORT is one of the huge estimations which is astoundingly speedy as well as accu-rate [10].

III. METHODOLOGY

The undertaking's work process mostly comprises of three stages, for example location of an item, following of article and including of the article in the zone. For the main stage, for example detec-tion of an article, we utilized YOLOv4 [11] with pre-prepared loads. YOLOv4 is an expansion in YOLOv3, with better speed and precision Article recognition models are prepared to characterize the items in a picture. . The principle justification utilizing YOLOv4 in our task is that it is effective for continuous article discovery, which is fundamental for our venture. The upside of constant recognition models is that they are not difficult to employ by all designers. YOLOv4 focuses on it for ongoing article location. Moreo-ver, the creators' planned for PC vision architects and designers to utilize their YOLOv4 structure in custom spaces without any problem.

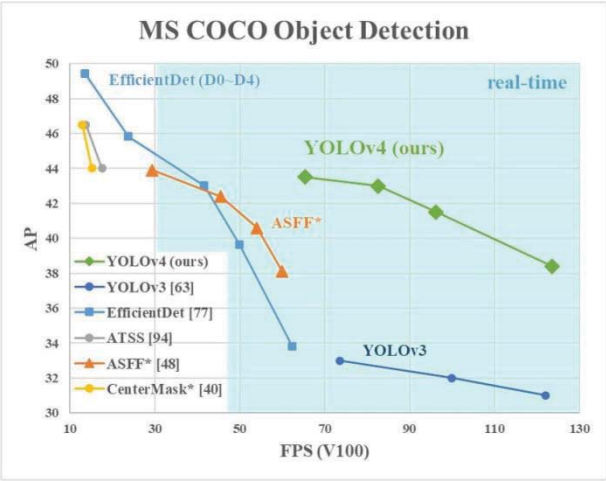


Fig. 1. Other detectors comparison vs YOLO

For the subsequent stage, for example following of items, we utilized the Deep SORT calculation. It is an augmentation to the SORT which is Simple Real-Time Tracker calculation. The DeepSORT calculation utilizes the kalman channels and it is necessary segment to gauge the roads of existing articles in the current edge. It has expressed that it contains eight factors; (u', u, v, 'v, a', a, h', h) where:

- (u,v) – focal points of the boxes bouncing
- h – Image height
- a – perspective extent
- variable other than this are specific there velocity

The Deep SORT uses a general and clear sort of Kalman channels which uses consistent rates as well as straight discernments. Exactly when each new edge comes, the circumstance of each and every track is surveyed dependent on its past regions. It isolates features so much that the features with different characters are at a critical distance isolated, and the features with a comparative character are almost each other. The new ID outcomes are connected with each incoming edges' present tracks utilizing a appearance incorporate vector and the track expected position. Simply the space info is used for track appraisal. By then we used the appearance incorporate vector to portray all of the features of a given picture. This vector gains the appearance info of recognizable pieces of proof by removing features and following pictures from past diagrams. Recognizable pieces of proof with low conviction regard than the edge are filtered through.

The vector is convolution neural association arranged on an enormous re-conspicuous verification dataset. New detec-tions need to pass this breaking point to transform into the contenders of data association. For the last development, which is checking the things, we have counted the amount of articles with the fascinating ID for calculating the amount of things in the zone. By then a cutoff is set for the zone. From the outset, the area is of green tone, and in case the amount of articles outperforms the cutoff, the alarm is ringed and the shade of the area changes to red.

IV. EXPERIMENTS AND RESULTS

This is a tremendous dataset con-taining unique, containing 16658 edges, 2412 tracks and 572834 bouncing boxes. We utilized the MOT18 dataset [12] for assessing the model.

Assessing Multiple Tracking of Objects messes some up as datasets are challeng-ing to build up. Different Object Tracking (MOT) calculations are subject to the indicator. The analysts don't generally concur on the convention to follow, i.e., regardless of whether to assess the tracker with a similar locator or assess the entire framework (indicator in addition to tracker) together? Albeit some presentation measures, for example, MOTA (Multi-ple Object Tracking Accuracy) are ordinarily utilized, a few analysts actually don't cling to this measurement.

In this, we have utilized Multiple Object Tracking Accuracy which is MOTA for assessing our results. MOTA is the most utilized rundown number in MOT. It can be characterized as:

Here we have ,

- FPt = ghost trajectories (No. of positive which are false)
- IDSt = At time t, identity switches number
- FNt = missed targets (No. of negatives which are false)
- GTt = truth of ground

Note: If ground truth is less than threshold, a target is assumed missed

TABLE I. MOTA SCORES

Sequence	Avg. Density	MOTA	FPS
MOT17-03	69.8	61.03%	12.61
MOT17-04	45.3	60.17%	19.43

We surveyed the model using the MOT18 dataset . It was done using Jetson Xavier NX. While using open area to use

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:50:46 UTC from IEEE Xplore. Restrictions apply.

the MOT18 dataset, the MOTA scores are close to bleeding edge trackers. Following rate can be reached upto 31 FPS (traces each second of time) dependent upon the article thickness. On a work territory FPS , GPU/CPU can be significantly higher. Regardless of the way that the tracker runs significantly faster, especially ac-priest. Lightweight finders, for instance, YOLOv4-little can be used to get greater fps up to 45(fps).

V. DISCUSSIONS

Following works so much that when it separates a specific thing, it will relegate a particu-lar exceptional id number to that specific article, and consequently, this intriguing character will change into the fundamental focal point for observing that specific article. Assigning out the remarkable IDs makes it all the more clear for the tracker to screen protests all through the video or imaging gathering (see Fig. 2).



Fig. 2. Object detection having distinguishable ids

During the article disclosure and following Process, there may happen a situation when a specific article is impeded by another article, which may really be the inspiration driving the assessment and can cause the thing in the assessment to vanish from the view totally or fairly (see Fig. 3). By applying the Kalman channels, this issue of hindrance is re-handled.

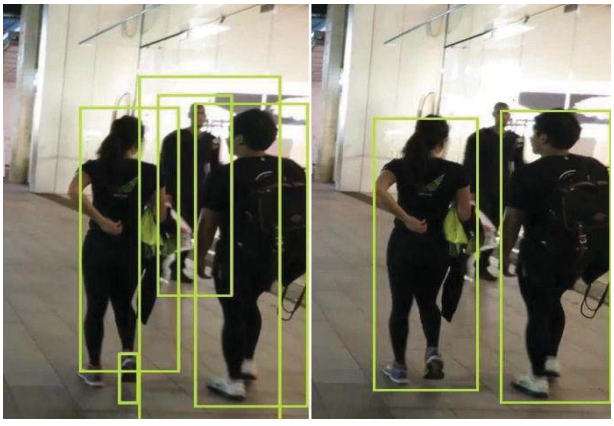


Fig. 3. Detections

The balance of the spread of Covid by placing this device in the Coronavirus space of interest domains and setting up a couple of zones, and raising alarms when the amount of people in the zone gets more unmistakable than the breaking point regard is one of the essential ap-plications of this article disclosure and worldwide situating structure in a specific zone. Exactly if the amount of person in the area is under 5 (edge regard), the area gets green (see Fig. 4), and if the amount of persons is more conspicuous than or identical to 5, by then it be-comes red (see Fig. 5).

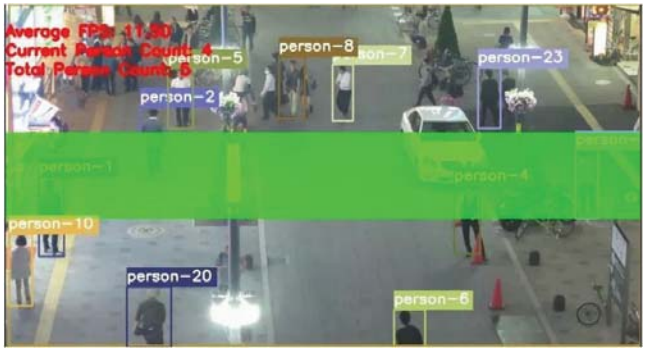


Fig. 4. Object count less than 5 makes zone green

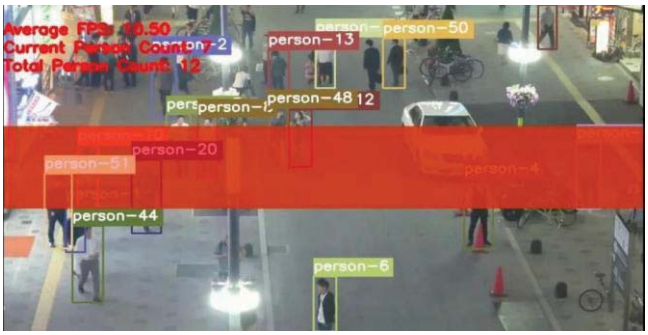


Fig. 5. Objects count greater than 5 make zone red

VI. CONCLUSION

With this, we are going towards the completion of the paper. Besides, it will in general be used to check the portion subtleties and see the quantity of peo-ple are moving beginning with one spot then onto the following, and through this data, govt. Ensuing to perceiving all of the articles in a given packaging, we will follow them until they are in the edge.

The experiment can have many businesses. For example, Tracing of vehicles can have many uses, for instance, checking the quantity of people are there in the leaving of a space. For significant and improved result, we have made use of kalman filters, and a previous SORT estimation was used, which addresses SORT ie simple online realtime tracking, but at this point we have used Deep SORT, which is a SORT computation with a deep connection figures. At last, we made use of the MOT-A figure as well as had the choice to make approx 60% accu-disgusting on area.

To summarize what we have done in this paper, we should rehash the troublesome we are overseeing: object recognizing verification and Tracking in a specific region where we will at first separate the thing using the YOLO show and see something comparative. Can give better inspirations and organizations to such people, and the public authority can similarly isolate the spots of interest and fix the security.

At long last, if there should arise an occurrence of any burglary in a specific district, it will in general be used to track, and we can alert the different experts included with the objective that the case can be dealt with even more quickly and without any problem. It will in general be used to screen

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:50:46 UTC from IEEE Xplore. Restrictions apply.

encroachment of travel guidelines, for instance, way advancing, speeding, and rash driving. It can in like manner be used to decrease time-handling gridlocks.

Because of specific disadvantages, for instance, awful environment, ill-advised lighting conditions during the evening, and less dataset availability, we couldn't achieve unprecedented results. Regardless, for extra degree, we can regardless improve the precision of the model by means of train-ing the model on a greater dataset and from this time forward extending the efficiency of the experiment. For further references, we should use the data which is stored for data assessment and make use of it to outfit the notice ministration with much better outcomes for figuring the leave plan in the era of crisis.

REFERENCES

[1] G. Cheng, J. Han, and X. Lu, “Remote sensing image scene classification: Benchmark and state of the art,” Proceedings of the IEEE, vol. 105, no. 10, pp. 1865–1883, 2017.

[2] S. Shinde, A. Kothari and V. Gupta, "YOLO based Human Action Recognition and Localization," Procedia Computer Science, vol. 133, pp. 831-838, 2018

[3] S. Kanimozhi, G. Gayathri, and T. Mala, "Multiple Real-time object identification using Single shot Multi-Box detection," in 2019 International Conference on Computational In- telligence in Data Science (ICCIDS), (2019).

[4] G. Swalaganata, Mun-iri and Y. Affriyenni, "Moving obje-ct tracking using hybrid method," 2018 International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, 2018, pp. 607-611, (2018).

[5]] D. Wen, X. Huang, L. Zhang, and J. A. Benediktsson, “A novel automatic change detection method for urban high-resolution remotely sensed imagery based on multiindex scene representation,” IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 1, pp. 609–625, 2015

[6] P. SHARMA, "A Step-by-Step Introduction to the Basic Object Detection Algorithms (Part 1)," Analytics Vidhya, 2019. [Online]. Available: https://www.analyticsvidhya.com/blog/2018/10/a-step-bystep-introduction-to-the-basic-object-detection-algorithms-part-1/.

[7] Dumitru Erhan, Christian Szegedy, Alexander Toshev, “Scalable Object Detection using Deep Neural Networks”, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 2147-2154.

[8] I. J. Goodfellow, J. Shlens, and C. Szegedy, “Explaining and harnessing adversarial examples,” 2014.Karen Simonyan, Andrew Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition”, published in Computer Vision and Pattern Recognition (cs.CV).

[9] N. Wojke, A. Bewley, and D. Paulus, "Simple online and realtime tracking with a deep as- sociation metric," in 2017 IEEE International Conference on Image Processing (ICIP), (2017).

[10] YOLO Juan Du1,”Understanding of Object Detection Based on CNN Family”,New Research, and Development Center of Hisense, Qingdao 266071, China.

[11] Y. Hu, J. Chen, D. Pan and Z. Hao, "Edge-Guided Image Object Detection in Multiscale Segmentation for High-Resolution Remotely Sensed Imagery," in IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 8, pp. 4702-4711, Aug. (2016).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:50:46 UTC from IEEE Xplore. Restrictions apply.

arXiv:2106.04962v1 [math.CV] 9 Jun 2021

ON CERTAIN GENERALIZATIONS OF $\mathcal{S}^*(\psi)$: II

KAMALJEET GANGANIA AND S. SIVAPRASAD KUMAR

Abstract. In this paper, we prove various radius results and obtain sufficient conditions using the convolution for the Ma-Minda classes $\mathcal{S}^*(\psi)$ and $\mathcal{C}(\psi)$ of starlike and convex analytic functions. We also obtain the Bohr radius for the class $S_f(\psi) := \{g(z) = \sum_{k=1}^{\infty} b_k z^k : g \prec f\}$ of subordinants, where $f \in \mathcal{S}^*(\psi)$. The results are improvements and generalizations of several well known results.

2010 AMS Subject Classification. Primary 30C45, 30C80, Secondary 3050.
Keywords and Phrases. Subordination, Bohr radius, Convolution, Distortion theorem.

1. Introduction

Let $\mathcal{A}$ denote the class of analytic functions of the form $f(z) = z + \sum_{k=2}^{\infty} a_k z^k$ in the open unit disk $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$. Using subordination, Ma-Minda [13] introduced the unified class of starlike and convex functions defined as follows:

$$\mathcal{S}^*(\psi) := \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} \prec \psi(z) \right\} \tag{1.1}$$

and

$$\mathcal{C}(\psi) := \left\{ f \in \mathcal{A} : 1 + \frac{zf''(z)}{f'(z)} \prec \psi(z) \right\},$$

where ψ is a Ma-Minda function, which is analytic and univalent with $\Re\psi(z) > 0$, $\psi'(0) > 0$, $\psi(0) = 1$ and $\psi(\mathbb{D})$ is symmetric about real axis. Note that $\psi \in \mathcal{P}$, the class of normalized Carathéodory functions. The class $\mathcal{S}^*(\psi)$ unifies various subclasses of starlike functions, which are obtained for an appropriate choice of ψ. See [6, 8, 10, 14, 18, 23]. Ma-Minda discussed many properties of the class $\mathcal{S}^*(\psi)$, in particular, they proved the distortion theorem [13, Theorem 2, p.162] with some restriction on ψ, namely

$$\min_{|z|=r} |\psi(z)| = \psi(-r) \quad \text{and} \quad \max_{|z|=r} |\psi(z)| = \psi(r). \tag{1.2}$$

Importance of the above conditions can be seen in achieving the sharpness for majorization results in [7]. In section 3, we modify the distortion theorem by relaxing this restriction on ψ to obtain a more general result. In

K. Gangania thanks to University Grant Commission, New-Delhi, India for providing Junior Research Fellowship under UGC-Ref. No.:1051/(CSIR-UGC NET JUNE 2017).

1914, Harald Bohr [4] proved the following remarkable result related to the power series:

Theorem 1.1 ([4]). *Let $g(z) = \sum_{k=0}^{\infty} a_k z^k$ be an analytic function in $\mathbb{D}$ and $|g(z)| < 1$ for all $z \in \mathbb{D}$, then $\sum_{k=0}^{\infty} |a_k| r^k \leq 1$ for all $z \in \mathbb{D}$ with $|z| = r \leq 1/3$.*

Bohr actually proved the above result for $r \leq 1/6$. Further Wiener, Riesz and Shur independently sharpened the result for $r \leq 1/3$. Presently the Bohr inequality for functions mapping unit disk onto different domains, other than unit disk is an active area of research. For the recent development on Bohr-phenomenon, see the articles [1, 2, 3, 15, 16] and references therein. The concept of Bohr-phenomenon in terms of subordination can be described as: Let $f(z) = \sum_{k=0}^{\infty} a_k z^k$ and $g(z) = \sum_{k=0}^{\infty} b_k z^k$ are analytic in $\mathbb{D}$ and $f(\mathbb{D}) = \Omega$. For a fixed f, consider a class of analytic functions $S(f) := \{g : g \prec f\}$ or equivalently $S(\Omega) := \{g : g(z) \in \Omega\}$. Then the class $S(f)$ is said to satisfy Bohr-phenomenon, if there exists a constant $r_0 \in (0, 1]$ satisfying the inequality $\sum_{k=1}^{\infty} |b_k| r^k \leq d(f(0), \partial\Omega)$ for all $|z| = r \leq r_0$ and $g(z) \in S(f)$, where $d(f(0), \partial\Omega)$ denotes the Euclidean distance between $f(0)$ and the boundary of $\Omega = f(\mathbb{D})$. The largest such r_0 for which the inequality holds, is called the Bohr-radius. In 2014, Muhanna et al. [15] proved the Bohr-phenomenon for $S(W_\alpha)$, where $W_\alpha := \{w \in \mathbb{C} : |\arg w| < \alpha\pi/2, 1 \leq \alpha \leq 2\}$, which is a Concave-wedge domain (or exterior of a compact convex set) and the class $R(\alpha, \beta, h)$ defined by $R(\alpha, \beta, h) := \{g \in \mathcal{A} : g(z) + \alpha z g'(z) + \beta z^2 g''(z) \prec h(z)\}$, where h is a convex function (or starlike) and $R(\alpha, \beta, h) \subset S(h)$. In 2018, Bhowmik and Das [2] proved the Bohr-phenomenon for the classes given by $S(f) = \{g \in \mathcal{A} : g \prec f \text{ and } f \in \mu(\lambda)\}$, where $\mu(\lambda) = \{f \in \mathcal{A} : |(z/f(z))^2 f'(z) - 1| < \lambda, 0 < \lambda \leq 1\}$ and $S(f) = \{g \in \mathcal{A} : g \prec f \text{ and } f \in \mathcal{S}^*(\alpha), 0 \leq \alpha \leq 1/2\}$, where $\mathcal{S}^*(\alpha)$ is the well-known class of starlike functions of order α.

In Section 4, for any fixed $f \in \mathcal{S}^*(\psi)$, we introduce and study the Bohr-phenomenon inside the disk $|z| \leq 1/3$ for the following class of analytic subordinants:

Definition 1.2. Let $f \in \mathcal{S}^*(\psi)$. Then the class of analytic subordinants functions is defined as

$$S_f(\psi) := \left\{ g(z) = \sum_{k=1}^{\infty} b_k z^k : g \prec f \right\}. \tag{1.3}$$

Note that $\mathcal{S}^*(\psi) \subset \bigcup_{f \in \mathcal{S}^*(\psi)} S_f(\psi)$. As an application, we obtain the Bohr-radius for the class $S(f)$, where $f \in \mathcal{S}^*((1 + Dz)/(1 + Ez))$, the class of Janowski starlike functions, with some additional restriction on D and E apart from $-1 \leq E < D \leq 1$.

Now recall that the convolution of two power series $f(z) = z+\sum_{n=2}^{\infty} a_n z^n$ and $g(z) = z + \sum_{n=2}^{\infty} b_n z^n$ is defined by

$$(f * g)(z) = z + \sum_{n=2}^{\infty} a_n b_n z^n.$$

Ruscheweyh and Sheil-Small [19] proved that if $f \in \mathcal{C}$ and $g \in \mathcal{C}$ ($g \in \mathcal{S}^*$), then $f * g \in \mathcal{C}$ ($f * g \in \mathcal{S}^*$). Later on Ma-Minda [13] proved that $f \in \mathcal{C}$ and $g \in \mathcal{S}^*(\psi)$, then $f * g \in \mathcal{S}^*(\psi)$, when $\psi(\mathbb{D})$ is convex. Szegö [24] in 1928 dicussed the radii of convexity for the sections $f_k(z) = z + \sum_{n=2}^{k} a_n z^n$ of the functions $f \in \mathcal{C}$, while in 1988 Silverman [22] considered the radii of starlikeness of f_k. Also see Silverman et al. [20, 21] work on convolution for the Janowski classes. We shall consider these problems for the Ma-Minda classes $\mathcal{C}(\psi)$ and $\mathcal{S}^*(\psi)$ following the idea of Goodman and Schoenberg [25]. In section 2, we obtain the condition for functions to be in $\mathcal{S}^*(\psi)$ which is an extention of the Bulboacă and Tuneski [5]. We also obtain several radius results and, necessary and sufficient conditions through the convolution for the classes $\mathcal{S}^*(\psi)$ and $\mathcal{C}(\psi)$.

2. Convolution properties and sufficient conditions for starlike and convex functions

In the following, we extend the results of the Bulboacă and Tuneski [5] for the class $\mathcal{S}^*(\psi)$:

Theorem 2.1. *Let h be analytic with $h(0) = 0$, $h'(0) \neq 0$. Suppose that h satisfies*

$$\Re\left(1 + \frac{zh''(z)}{h(z)}\right) > -\frac{1}{2} \tag{2.1}$$

and

$$\frac{1}{z}\int_0^z h(t)dt \prec \frac{\psi(z) - 1}{\psi(z)}. \tag{2.2}$$

If $f \in \mathcal{A}$, then

$$\frac{f(z)f''(z)}{(f'(z))^2} \prec h(z) \quad \textit{implies} \quad f \in \mathcal{S}^*(\psi).$$

Proof. Using the result [5, Theorem 3.1, p. 3], we see that $f(z)f''(z)/(f'(z))^2 \prec h(z)$ implies

$$\frac{1}{z}\int_0^z \left(1 - \left(\frac{f(t)}{f'(t)}\right)'\right) dt = 1 - \frac{f(z)}{zf'(z)} \prec \frac{1}{z}\int_0^z h(t)dt.$$

From the above subordination, we have

$$\frac{f(z)}{zf'(z)} \prec 1 - \frac{1}{z}\int_0^z h(t)dt.$$

Now to prove that $f \in \mathcal{S}^*(\psi)$, it suffices to consider

$$1 - \frac{1}{z}\int_0^z h(t)dt \prec \frac{1}{\psi(z)},$$

which is equivalent to (2.2). This completes the proof. □

Corollary 2.2. *Let* $-1 \leq E < D \leq 1$ *and* $0 < \beta \leq 1$ *such that*

$$h(z) = 1 - \left(\frac{1+Ez}{1+Dz}\right)^{\beta} \left(\frac{1+(D+E-\beta(D-E))z+DEz^2}{(1+Dz)(1+Ez)}\right)$$

and satisfies (2.1). *If* $f \in \mathcal{A}$ *and*

$$\frac{f(z)f''(z)}{(f'(z))^2} \prec h(z).$$

Then $f \in \mathcal{S}^*\left(\left(\frac{1+Dz}{1+Ez}\right)^{\beta}\right)$.

Remark 2.1. (*i*) Choosing $E = -1$ and $D = 1$ in Corollary 2.2 gives [5, Corollary 3.5].
(*ii*) Choosing $E = 0$ and $\beta = 1$, then Corollary 2.2 reduces to [5, Example 4.5].

Choosing $D = 1 - 2\alpha$, $E = -1$ and $\beta = 1$ in Corollary 2.2 yields [5, Corollary 3.4] (Note that $\alpha \neq 1$). Also, in [5, Example 4.4] correct range of α is $[0, 1)$.

Corollary 2.3. *If* $f \in \mathcal{A}$, $0 \leq \alpha < 1$ *and*

$$\frac{f(z)f''(z)}{(f'(z))^2} \prec \frac{2(1-\alpha)((1-2\alpha)z^2+2z)}{(1+(1-2\alpha)z)^2}.$$

Then f *is starlike of order* α.

Considering $\psi(z) = \sqrt{1+cz}$, where $0 < c \leq 1$ in Theorem 2.1, we get the following:

Corollary 2.4. *Let* $0 < c < c_0 < 1$, *where* $c_0 \approx 0.845276$ *is the unique positive root of the equation*

$$30 - \frac{75}{2}c^2 - \frac{201}{32}c^4 = 0.$$

If $f \in \mathcal{A}$ *and*

$$\frac{f(z)f''(z)}{(f'(z))^2} \prec 1 - \frac{2+cz}{2(1+cz)^{3/2}}.$$

Then $f \in \mathcal{S}^*(\sqrt{1+cz})$.

Let us consider the following function in Theorem 2.1:

$$h(z) = 1 - \frac{e^{\lambda z}(1-\lambda z)}{e^{2\lambda z}},$$

where $0 < \lambda \leq 1$. Then

$$1 + \frac{zh''(z)}{h'(z)} = \frac{2-4\lambda z+\lambda^2 z^2}{2-\lambda z} \quad \text{and} \quad \Re\left(1+\frac{zh''(z)}{h'(z)}\right) = \frac{2-4\lambda+\lambda^2}{2-\lambda}.$$

Corollary 2.5. *Let* $0 < \lambda \leq \frac{1}{4}(9 - \sqrt{33}) < 1$. *If* $f \in \mathcal{A}$ *and*

$$\frac{f(z)f''(z)}{(f'(z))^2} \prec 1 - \frac{e^{\lambda z}(1-\lambda z)}{e^{2\lambda z}}.$$

Then $f \in \mathcal{S}^*(e^{\lambda z})$.

In 1985, Silverman and Silvia [21] obtained some necessary and sufficient conditions in terms of convolution operators for the functions to be in the Janowski classes, and generalized the results of Silverman et. al. [20]. We now extend the results for classes $\mathcal{S}^*(\psi)$ and $\mathcal{C}(\psi)$ as follows:

Theorem 2.6. $f \in \mathcal{S}^*(\psi)$ *if and only if*

$$\frac{1}{z}\left(f(z) * \frac{z - \lambda z^2}{(1-z)^2}\right) \neq 0, \tag{2.3}$$

where $\lambda = \psi(e^{it})/(1 - \psi(e^{it}))$ *and* $t \in [0, 2\pi)$.

Proof. $f \in \mathcal{S}^*(\psi)$ if and only if $zf'(z)/f(z) \prec \psi(z)$, which is further equivalent to

$$\frac{zf'(z)}{f(z)} \neq \psi(e^{it}), \quad (z \in \mathbb{D}, t \in [0, 2\pi)),$$

(means that the values $zf'(z)/f(z)$ does not lie on the boundary Ω of the domain $\psi(\mathbb{D})$) which, using the fact that $f(z)/z \neq 0$, can be equivalently written as

$$\frac{1}{z}\left(zf'(z) - \psi(e^{it})f(z)\right) \neq 0. \tag{2.4}$$

Since

$$zf'(z) = f(z) * \frac{z}{(1-z)^2} \quad \text{and} \quad f(z) = f(z) * \frac{z}{1-z}.$$

Therefore, (2.4) becomes

$$\frac{1}{z}\left(f(z) * \frac{z - \psi(e^{it})(z - z^2)}{(1-z)^2}\right) \neq 0,$$

and further with a little computation, it reduces to (2.3). □

Note that we can also write

$$\frac{1}{z}\left(f(z) * \frac{z - \lambda z^2}{(1-z)^2}\right) = \frac{1}{z}(zf'(z) - \lambda(zf'(z) - f(z))).$$

Thus using the power series expansion of $f(z)$, (2.3) becomes

$$\sum_{k=2}^{\infty}(\lambda(k-1) - k)a_k z^{k-1} \neq 1,$$

where λ is as defined in Theorem 2.6, and yields the following sufficient condition in terms of coefficients for the functions from $\mathcal{A}$ to be in $\mathcal{S}^*(\psi)$:

Corollary 2.7. *If* $f \in \mathcal{A}$ *satisfies* $\sum_{k=2}^{\infty} |\lambda(k-1) - k||a_k| < 1$. *Then* $f \in \mathcal{S}^*(\psi)$.

In particular, we have the following sufficient condition for the class of Janowski starlike functions:

Corollary 2.8. *The function $f \in \mathcal{S}^*((1+Dz)/(1+Ez))$, if*

$$\sum_{k=2}^{\infty}\left(\frac{1+|D|}{D-E}+\left(\frac{1+|D|-E+D}{D-E}\right)k\right)|a_k|<1.$$

The following result with some rearrangment and specific values of D and E reduces to [21, Theorem 6, Theorem 7].

Corollary 2.9. *Let $\psi(z)=(1+Dz)/(1+Ez)$, where $-1\leq E<D\leq 1$. Then $f\in\mathcal{S}^*(\psi)$ if and only if*

$$\frac{1}{z}\left(f(z)*\frac{z+\frac{\zeta+D}{D-E}z^2}{(1-z)^2}\right)\neq 0,\quad |\zeta|=1.$$

Corollary 2.10. *Let $\psi(z)=1+ze^z$. Then $f\in\mathcal{S}^*(\psi)$ if and only if*

$$\frac{1}{z}\left(f(z)*\frac{z+(1+\zeta e^{\zeta})z^2}{(1-z)^2}\right)\neq 0,\quad |\zeta|=1.$$

The class $\mathcal{S}^*(1+\sin z)$ was introduced in [6].

Corollary 2.11. *Let $\psi(z)=1+\sin z$. Then $f\in\mathcal{S}^*(\psi)$ if and only if*

$$\frac{1}{z}\left(f(z)*\frac{z-\arcsin\zeta(1+\sin\zeta)z^2}{(1-z)^2}\right)\neq 0,\quad |\zeta|=1.$$

Note that $f\in\mathcal{S}^*(\psi)$ if and only if $g(z)=\int_0^z\frac{f(t)}{t}dt\in\mathcal{C}(\psi)$. Therefore, the condition given in (2.3) is equivalent to the following:

$$\frac{1}{z}\left(zg'(z)*\frac{z-\lambda z^2}{(1-z)^2}\right). \tag{2.5}$$

Now using the convolution fact that $zg'(z)*f(z)=g(z)*zf'(z)$ in (2.5), we obtain the following result, which is the convex analogue of Theorem 2.6:

Theorem 2.12. *$f\in\mathcal{C}(\psi)$ if and only if*

$$\frac{1}{z}\left(f(z)*\frac{z+(1-2\lambda)z^2}{(1-z)^3}\right)\neq 0,$$

where λ is as defined in Theorem 2.6.

Note that we can also write

$$\frac{1}{z}\left(f(z)*\frac{z+(1-2\lambda)z^2}{(1-z)^3}\right)=f'(z)+(1-\lambda)zf''(z).$$

Thus, using the power series expansion of $f(z)$, we equivalently get

$$\sum_{k=2}^{\infty}(\lambda(k-1)-k)ka_kz^{k-1}\neq 1,$$

which gives the following sufficient condition in terms of coefficient's:

Corollary 2.13. *If $f \in \mathcal{A}$ satisfies $\sum_{k=2}^{\infty} k|\lambda(k-1)-k||a_k| < 1$. Then $f \in \mathcal{C}(\psi)$.*

In particular, we have the following sufficient condition for the class of Janowski convex functions:

Corollary 2.14. *The function $f \in \mathcal{C}((1+Dz)/(1+Ez))$, if*

$$\sum_{k=2}^{\infty} \left(\frac{1+|D|}{D-E} + \left(\frac{1+|D|-E+D}{D-E} \right) k \right) k|a_k| < 1.$$

Corollary 2.15. *Let $\psi(z) = (1+Dz)/(1+Ez)$, where $-1 \leq E < D \leq 1$. Then $f \in \mathcal{C}(\psi)$ if and only if*

$$\frac{1}{z}\left(f(z) * \frac{z + \frac{(3D-E)+2\zeta}{D-E} z^2}{(1-z)^3} \right) \neq 0, \quad |\zeta| = 1.$$

Corollary 2.16. *Let $\psi(z) = 1 + ze^z$. Then $f \in \mathcal{C}(\psi)$ if and only if*

$$\frac{1}{z}\left(f(z) * \frac{z + (3 + 2\zeta e^{\zeta}) z^2}{(1-z)^3} \right) \neq 0, \quad |\zeta| = 1.$$

Corollary 2.17. *Let $\psi(z) = 1 + \sin z$. Then $f \in \mathcal{C}(\psi)$ if and only if*

$$\frac{1}{z}\left(f(z) * \frac{z - \arcsin\zeta(2 + \sin\zeta) z^2}{(1-z)^3} \right) \neq 0, \quad |\zeta| = 1.$$

Remark 2.2. Let us consider the class introduced in [11]:

$$\mathcal{F}(q) := \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} \prec q(z) \right\},$$

where $q(z) = 1 + \sum_{n=2}^{\infty} q_n z^n$ is univalent. Further, assume that $f(z)/z \neq 0$. Then Theorem 2.6 also holds for the class $\mathcal{F}(q)$, and Theorem 2.12 for its convex analogue.

It is well known that the class $\mathcal{C}(\psi)$ is closed under convolution. Also the class $\mathcal{S}^*(\psi)$ is closed under convolution when convoluted with convex functions, when $\psi(\mathbb{D})$ is convex. With this motivation, the following result provides the largest radius r_0 such that in $|z| < r_0$ the class $\mathcal{S}^*(\psi)$ is closed under convolution.

Theorem 2.18. *Let r_ψ be the largest radius such that $F(z) = z + \sum_{n=2}^{\infty} n^2 z^n$ belongs to $\mathcal{S}^*(\psi)$ for $|z| < r_\psi$. If $f, g \in \mathcal{S}^*(\psi)$, where ψ is convex. Then $f * g$ belongs to $\mathcal{S}^*(\psi)$ for $|z| < r_\psi$. The radius is best possible.*

Proof. Let $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$ and $g(z) = z + \sum_{n=2}^{\infty} b_n z^n$. Then

$$\begin{aligned} G(z) &:= f(z) * g(z) \\ &= \left(z + \sum_{n=2}^{\infty} n^2 z^n \right) * \left(z + \sum_{n=2}^{\infty} \frac{a_n}{n} z^n \right) * \left(z + \sum_{n=2}^{\infty} \frac{b_n}{n} z^n \right) \end{aligned}$$

$$= F(z) * \left(z + \sum_{n=2}^{\infty} \frac{a_n}{n} z^n \right) * \left(z + \sum_{n=2}^{\infty} \frac{b_n}{n} z^n \right).$$

Recall that

$$f \in \mathcal{S}^*(\psi) \quad \text{if and only if} \quad \int_0^z \frac{f(t)}{t} dt \in \mathcal{C}(\psi).$$

Similary for the function g. Therefore, $z + \sum_{n=2}^{\infty} \frac{a_n}{n} z^n$ and $z + \sum_{n=2}^{\infty} \frac{b_n}{n} z^n$ belong to $\mathcal{C}(\psi)$. Now let

$$\begin{aligned} H(z) &:= \left(z + \sum_{n=2}^{\infty} \frac{a_n}{n} z^n \right) * \left(z + \sum_{n=2}^{\infty} \frac{b_n}{n} z^n \right) \\ &= z + \sum_{n=2}^{\infty} \frac{a_n b_n}{n^2} z^n \in \mathcal{C}(\psi) \subseteq \mathcal{C} \end{aligned}$$

so that

$$G(z) = F(z) * H(z).$$

Now from the hypothesis, we have

$$\frac{F(r_\psi z)}{r_\psi} \in \mathcal{S}^*(\psi), \tag{2.6}$$

that is, F belongs to $\mathcal{S}^*(\psi)$ for $|z| < r_\psi$. Since $H \in \mathcal{C}$. Therefore, using (2.6), we get

$$B(z) := \left(\frac{F(r_\psi z)}{r_\psi} \right) * H(z) \in \mathcal{S}^*(\psi), \tag{2.7}$$

whenever ψ is convex. Thus, we conclude that

$$G(z) = r_\psi B\left(\frac{z}{r_\psi} \right) \in \mathcal{S}^*(\psi), \quad (|z| < r_\psi).$$

Note that the radius r_ψ is independent of the choices of the functions f and g, and thus the sharpness of the result follows from (2.6). □

Remark 2.3. Note that

$$F(z) = \frac{z(1+z)}{(1-z)^3} = z + \sum_{n=2}^{\infty} n^2 z^n.$$

Now the logarithmic differentiation of the above yields:

$$\frac{zF'(z)}{F(z)} = \frac{1 + 4z + z^2}{1 - z^2} = \frac{1+z}{1-z} - \frac{1}{1+z} + \frac{1}{1-z}. \tag{2.8}$$

It follows from (2.8) that

$$\Re\left(\frac{zF'(z)}{F(z)} \right) \geq \frac{1 - 4r + r^2}{1 - r^2}. \tag{2.9}$$

Moreover, the following sharp inequality also holds:

$$\left| \frac{zF'(z)}{F(z)} - 1 \right| \leq \frac{2r(2+r)}{1 - r^2}. \tag{2.10}$$

Furthermore, the following inequality also holds:

$$\left|\frac{zF'(z)}{F(z)}-\frac{1+r^2}{1-r^2}\right|\leq\frac{4r}{1-r^2}. \tag{2.11}$$

Now applying Theorem 2.18, we get the following result:

Corollary 2.19. *Let $f,g\in\mathcal{S}^*(\psi)$, where ψ is convex. Then $f*g$ belongs to $\mathcal{S}^*(\psi)$ for $|z|<r_\psi$, where*

(i) $r_\psi=(2-\sqrt{3+\alpha^2})/(1+\alpha)$, *when* $\psi(z)=(1+(1-2\alpha)z)/(1-z)$.
(ii) $r_\psi=(-2+\sqrt{5})/(1+\sqrt{2})$, *when* $\psi(z)=\sqrt{1+z}$.
(iii) $r_\psi=(-2+\sqrt{4+b(2+b)})/(2+b)$, *where* $b=(e-1)/(e+1)$, *when* $\psi(z)=2/(1+e^{-z})$.
(iv) $r_\psi=(2-\sqrt{4-b^2})/b$, *where* $b=\sin\pi\gamma/2$, *when* $\psi(z)=((1+z)/(1-z))^\gamma$.

The radii are sharp.

Proof. The part (i) follows using the inequality (2.9) such that $(1-4r+r^2)(1-r^2)\geq\alpha$, which holds for $r\leq(2-\sqrt{3+\alpha^2})/(1+\alpha)$. Further

$$\frac{z_0F'(z_0)}{F(z_0)}=\alpha\quad\text{for}\quad z_0=\frac{\sqrt{3+\alpha^2}-2}{1+\alpha},$$

implies the sharpness of the radius. Note that the disks $\{w:|w-1|<\sqrt{2}-1\}$ and $\{w:|w-1|<(e-1)/(e+1)\}$ are contained in $\phi(\mathbb{D})$ where $\phi(z)=\sqrt{1+z}$ and $2/(1+e^{-z})$, respectively. Therefore, the parts (ii) and (iii) follow using the inequality (2.10) such that

$$\frac{2r(2+r)}{1-r^2}\leq\sqrt{2}-1\quad\text{and}\quad\frac{2r(2+r)}{1-r^2}\leq\frac{e-1}{e+1},$$

which holds for $r\leq(-2+\sqrt{5})/(1+\sqrt{2})$ and $r\leq(-2+\sqrt{4+b(2+b)})/(2+b)$, where $b=(e-1)/(e+1)$ respectively. Since

$$\frac{z_0F'(z_0)}{F(z_0)}=1+\sqrt{2}\quad\text{for}\quad z_0=\frac{-2+\sqrt{5}}{1+\sqrt{2}}$$

and

$$\frac{z_0F'(z_0)}{F(z_0)}=\frac{2}{1+e}\quad\text{for}\quad z_0=(\sqrt{4+b(2+b)}-2)/(2+b),$$

therefore the radii obtained are sharp. Part (iv) follows by using the inequality (2.11) and the fact that the disk $\{w:|w-a|<r_a\}$ is contained in the sector $|\arg w|\leq\pi\gamma/2$, whenever $r_a\leq a\sin(\pi\gamma/2)$. □

The following lemma was introduced by Kumar and Gangania [12] to obtain certain radius constants (see [12, p.12-14]) related to the operators like Livingston and Bernardi etc. to cover the case when ψ is starlike but not convex.

Lemma 2.1. *Let r_c be the radius of convexity of ψ. If $g\in\mathcal{C}$ and $f\in\mathcal{S}^*(\psi)$. Then $f*g\in\mathcal{S}^*(\psi)$ for $|z|<r_\psi=\min\{r_c,1\}$.*

Now using the above lemma, we may write (2.7) as follows:

$$B(z) := \left(\frac{F(r_0 z)}{r_0}\right) * H(z) \in \mathcal{S}^*(\psi),$$

where $r_0 = \min\{r_\psi, r_c\}$. Note that $r_c = (3-\sqrt{5})/2$ when $\psi(z) = 1 + ze^z$. Thus, we have the following result:

Corollary 2.20. *Let $f, g \in \mathcal{S}^*(1+ze^z)$. Then $f * g$ belongs to $\mathcal{S}^*(1+ze^z)$ for $|z| < (2e - \sqrt{4e^2 - 2e + 1})/(2e-1) \approx 0.0957$. The radius is sharp.*

Corollary 2.21. *Let $f, g \in \mathcal{S}^*(1+\sin z)$. Then $f * g$ belongs to $\mathcal{S}^*(1+\sin z)$ for $|z| < (\sqrt{4 + \sin 1(2+\sin 1)} - 2)/(2+\sin 1) \approx 0.1858$. The radius is sharp.*

In the next theorem, we continue to extend the ideas of Szegö [24] and Silverman [22] on the starlikeness and convexity of sections $f_k(z) = z + \sum_{n=2}^{k} a_n z^n$ of f in $\mathcal{S}^*(\psi)$ and $\mathcal{C}(\psi)$, respectively.

Theorem 2.22. *Let $g_k(z) = (z - z^{k+1})/(1-z)$. If $f \in \mathcal{C}(\psi)$, where ψ be convex. Then*

(i) $f_k \in \mathcal{C}(\psi)$ in $|z| < r_0$, where r_0 is the radius of convexity of g_k.
(ii) $f_k \in \mathcal{S}^*(\psi)$ in $|z| < r_\psi$, whenever g_k belongs to $\mathcal{S}^*(\psi)$ in $|z| < r_\psi$.

The radii are best possible.

Proof. Let $t = r_0$ in the first part and r_ψ in second part, respectively.Then the proof follows by observing that $f_k(z) = t h_k(z/t)$, where $h_k(z) = f(z) * \dfrac{g_k(tz)}{t}$. □

Remark 2.4. If we choose $\psi(z) = (1+z)/(1-z)$. Then Theorem 2.22-(ii) reduces to the Silverman's result [22, Theorem 1, p. 1192].

Jackson [9] introduced and studied the q-derivative defined as

$$d_q f(z) := \frac{f(qz) - f(z)}{(q-1)z} = \frac{1}{z}\left(z + \sum_{n=2}^{\infty} [n]_q a_n z^n\right), \quad z \neq 0$$

and $d_q f(0) = f'(0)$, where $[n]_q = \frac{1-q^n}{1-q}$.

Theorem 2.23. *Let r_ψ be the largest radius in $(0,1]$ and $q \in (0,1)$ such that*

$$\frac{z}{(1-qz)(1-z)} \in \mathcal{S}^*(\psi) \quad \text{for} \quad |z| < r_\psi,$$

where ψ is convex. If $f \in \mathcal{C}$, then we have

$$z d_q f(z) = z + \sum_{n=2}^{\infty} [n]_q a_n z^n \in \mathcal{S}^*(\psi) \quad \text{for} \quad |z| < r_\psi.$$

The radius is best possible.

Proof. Observe that for each $q \in \mathbb{C}$ where $|q| \leq 1$, $q \neq 1$, we have

$$h_q(z) = \frac{1}{1-q}\log\left(\frac{1-qz}{1-z}\right) = \sum_{n=1}^{\infty}\left(\frac{1-q^n}{1-q}\right)\frac{z^n}{n}$$
$$= \sum_{n=1}^{\infty}\frac{[n]_q}{n}z^n \in \mathcal{C},$$

which implies that

$$zh_q'(z) = \frac{z}{(1-qz)(1-z)} = \sum_{n=1}^{\infty}[n]_q z^n \in \mathcal{S}^*.$$

We note that

$$zd_qf(z) = \left(z + \sum_{n=2}^{\infty}a_nz^n\right) * \left(\sum_{n=1}^{\infty}[n]_qz^n\right)$$
$$= f(z) * \frac{z}{(1-qz)(1-z)}$$
$$= f(z) * zh_q'(z).$$

Now for simplicity, let us write $H(z) = zh_q'(z)$. Then from hypothesis we see that

$$\frac{H(r_\psi z)}{r_\psi} \in \mathcal{S}^*(\psi). \tag{2.12}$$

Since $f \in \mathcal{C}$, and $\mathcal{C} * \mathcal{S}^*(\psi) = \mathcal{S}^*(\psi)$ whenever ψ is convex. Therefore, we have

$$f(z) * \frac{H(r_\psi z)}{r_\psi} \in \mathcal{S}^*(\psi),$$

which is equivalent to saying that

$$zd_qf(z) \in \mathcal{S}^*(\psi) \quad \text{for} \quad |z| < r_\psi.$$

From the proof, we note that the sharpness of the radius r_ψ follows from (2.12). □

Remark 2.5. From the function $H(z) = z/((1-qz)(1-z))$, we have:

$$\frac{zH'(z)}{H(z)} = 1 + \frac{z}{1-z} + \frac{qz}{1-qz}. \tag{2.13}$$

It follows from (2.13) that

$$\Re\left(\frac{zH'(z)}{H(z)}\right) \geq \frac{1-qr^2}{(1+r)(1+qr)}. \tag{2.14}$$

Moreover, the following sharp inequality also holds:

$$\left|\frac{zH'(z)}{H(z)} - 1\right| \leq \frac{r(1+q-2qr)}{(1-r)(1-qr)}. \tag{2.15}$$

Now proceeding in a similar way as in Corollary 2.19 using (2.14) when $\psi(z) = (1+(1-2\alpha)z)/(1-z)$, and (2.15) when $\psi(z) = \sqrt{1+z}$ and $2/(1+e^{-z})$, we have

Corollary 2.24. *If $f \in \mathcal{C}$, then for all $0 < q < 1$, we have*

$$zd_q f(z) = z + \sum_{n=2}^{\infty} [n]_q a_n z^n \in \mathcal{S}^*(\psi) \quad for \quad |z| < r_\psi,$$

where

(i) $r_\psi = (\sqrt{\alpha^2(1-q^2)+4q} - \alpha(q+1))/(2q(1+\alpha))$, *when* $\psi(z) = (1+(1-2\alpha)z)/(1-z)$ *for* $\alpha \geq (1-q)/2(1+q)$.
(ii) $r_\psi = ((1+q) - \sqrt{1+q^2})/(q\sqrt{2}(\sqrt{2}+1))$, *when* $\psi(z) = \sqrt{1+z}$.
(iii) $r_\psi = ((1+q)(1+b) - \sqrt{((1+q)(1+b))^2 - 4bq(2+b)})/(2q(2+b))$, *where* $b = (e-1)/(e+1)$, *when* $\psi(z) = 2/(1+e^{-z})$.

The radii are sharp.

Remark 2.6. Note that part (i) of the above corollary includes the result [[17], Theorem 2.1]. Also $r_\psi = 1$ for $\alpha \in [0, (1-q)/2(1+q)]$.

3. **Distortion theorem**

Ma-Minda [13] proved the distortion theorem for the class $\mathcal{S}^*(\psi)$ with some restriction on ψ, namely $|\psi(z)|$ attains its maximum and minimum value respectively at $z = r$ and $z = -r$, see eq. (1.2). Now what if ψ does not satisfy the condition (1.2) and why the condition (1.2) is so important? To answer this, we first need to recall the following result:

Lemma 3.1. ([13]) *Let $f \in \mathcal{S}^*(\psi)$ and $|z_0| = r < 1$. Then $-f_0(-r) \leq |f(z_0)| \leq f_0(r)$. Equality holds for some $z_0 \neq 0$ if and only if f is a rotation of f_0, where $zf_0(z)/f_0(z) = \psi(z)$ such that*

$$f_0(z) = z \exp \int_0^z \frac{\psi(t) - 1}{t} dt. \tag{3.1}$$

We see that a Ma-Minda starlike function, in general, need not satisfy the condition (1.2). To examine the same, let us consider two different Ma-Minda starlike functions, namely $\psi_1(z) := z + \sqrt{1+z^2}$ and $\psi_2(z) := 1 + ze^z$. The unit disk images of ψ_1 and ψ_2 are displayed in figure 1 and figure 2.

We know that the radius of a circle, centered at origin and touching only the boundary points of a image domain of a complex function, yields the optimal values of the modulus of the function. For example, see figure 1 to locate the lower bound of the modulus for a crescent function. Therefore it is evident from figure 1 that both the bounds $\psi_1(-r)$ and $\psi_1(r)$ of $|\psi_1|$ are attained on the real line and we have $\psi_1(-r) \leq |\psi_1(z)| \leq \psi_1(r)$ for each $|z| = r$. Whereas, from figure 2, we see that only the upper bound $\psi_2(r)$ of $|\psi_2|$ is attained on the real line and $|\psi_2(z)| \leq \psi_2(r)$ for each $|z| = r$. Although both ψ_1 and ψ_2 are Ma-Minda functions but the distortion theorem of Ma-Minda [13, theorem 2, p. 162] does not accommodate the function ψ_2, as the lower bound for $|\psi_2(z)|$ is not attained on the real line for all $|z| = r > (3 - \sqrt{5})/2$, see figure 3. To overcome this limitation, we modify the distortion theorem, wherein we theoretically assume the

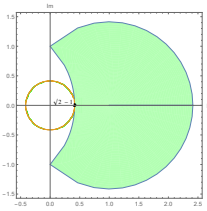


FIGURE 1. $\psi_1(z) := z + \sqrt{1+z^2}$

Let $\min_{|z|=r}|\psi_2(z)| =: \gamma_i(r)$, where $z = r_i e^{i\theta}$, $0 \le \theta \le \pi$, then from table 1, we have $\gamma_1(1) = 0.372412$, $\gamma_2(4/5) = 0.527912$, $\gamma_3(2/3) = 0.611553$, $\gamma_4(1/2) = 0.693287$, $\gamma_5(r) = 1 - re^{-r}$, where $r \le (3-\sqrt{5})/2$.

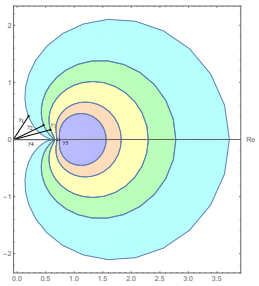


FIGURE 2. $\psi_2(z) := 1 + ze^z$

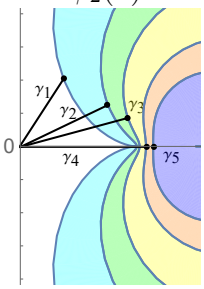


FIGURE 3. A zoom of figure 2

modulus bounds of the function and obtain a more general result. Thus the Ma-Minda functions, for which modulus bounds are not attained on the real line but could be computed can now be entertained for distortion theorem using the following result:

Theorem 3.1 (Modified Distortion Theorem)**.** *Let ψ be a Ma-Minda function. Assume that $\min_{|z|=r}|\psi(z)| = |\psi(z_1)|$ and $\max_{|z|=r}|\psi(z)| = |\psi(z_2)|$, where $z_1 = re^{i\theta_1}$ and $z_2 = re^{i\theta_2}$ for some $\theta_1, \theta_2 \in [0, \pi]$. Let $f \in \mathcal{S}^*(\psi)$ and $|z_0| = r < 1$. Then*

$$|\psi(z_1)|\left(\frac{-f_0(-r)}{r}\right) \le |f'(z_0)| \le \left(\frac{f_0(r)}{r}\right)|\psi(z_2)|. \tag{3.2}$$

Proof. Let $p(z) = zf'(z)/f(z)$. Then $f \in \mathcal{S}^*(\psi)$ if and only if $p(z) \prec \psi(z)$. Using a result [13, Theorem 1, p.161], we have

$$\frac{f(z)}{z} \prec \frac{f_0(z)}{z}, \tag{3.3}$$

where f_0 is given by (3.1). Now using Maximum-Minimum principle of modulus, (3.3) and by Lemma 3.1, $-f_0(-r)/r \leq |f(z_0)/z| \leq f_0(r)/r$, we easily obtain for $|z_0| = r$

$$\begin{aligned}
|\psi(z_1)| \left(\frac{-f_0(-r)}{r}\right) &= \min_{|z|=r} |\psi(z)| \min_{|z|=r} \left|\frac{f_0(z)}{z}\right| \\
&\leq \left|p(z_0)\frac{f(z_0)}{z_0}\right| = |f'(z_0)| \\
&\leq \max_{|z|=r} |\psi(z)| \max_{|z|=r} \left|\frac{f_0(z)}{z}\right| = \left(\frac{f_0(r)}{r}\right) |\psi(z_2)|,
\end{aligned}$$

that is,

$$|\psi(z_1)| \left(\frac{-f_0(-r)}{r}\right) \leq |f'(z_0)| \leq \left(\frac{f_0(r)}{r}\right) |\psi(z_2)|,$$

where z_1 and z_2 are as defined in the hypothesis. Hence the result. □

To illustrate Theorem 3.1, we consider the function $\psi(z) = 1 + ze^z$. Then we have the following expression for its modulus:

$$|\psi(z)| = \sqrt{1 + re^{r\cos\theta}(re^{r\cos\theta} + 2\cos(\theta + r\sin\theta))}. \tag{3.4}$$

Using equation (3.4) and Theorem 3.1, we obtain the following table, providing the minimum for various choices of r:

TABLE 1. The lower bounds for $|1 + ze^z|$ for different choices of $r = |z|$.

r	$0 \leq \theta_1 \leq \pi$	$\lvert\psi(re^{i\theta_1})\rvert$	$m(r,\theta_1) = \lvert\psi(re^{i\theta_1})\rvert(-f_0(-r)/r)$
1	1.88438	0.372412	0.197923
4/5	2.01859	0.527912	0.304374
2/3	2.17677	0.611553	0.375966
1/2	2.58169	0.693287	0.467769
$r \leq (3-\sqrt{5})/2$	π	$\psi_2(-r)$	$f_0'(-r)$

Now using Theorem 3.1, we obtain the following distortion theorem for the class $\mathcal{S}^*(1 + ze^z)$:

Example 3.2. *Let $\psi(z) = 1+ze^z$ and $\min_{|z|=r}|\psi(z)| = |\psi(z_1)|$, where $z_1 = re^{i\theta_1}$ for some $\theta_1 \in [0,\pi]$. Let $f \in \mathcal{S}^*(\psi)$ and $|z_0| = r < 1$. Then*

$$m(r,\theta_1) \leq |f'(z_0)| \leq f_0'(r), \quad \left(r > \tfrac{3-\sqrt{5}}{2}\right)$$

and

$$f_0'(-r) \leq |f'(z_0)| \leq f_0'(r), \quad \left(r \leq \tfrac{3-\sqrt{5}}{2}\right),$$

where $f_0(z) = z\exp(e^z - 1)$ *and* $m(r, \theta_1)$ *is provided in table 1 for some specific values of* r.

Remark 3.1. In Theorem 3.1, if we assume that $\theta_1 = \pi$ and $\theta_2 = 0$, then extremes in equation (3.2) simplifies to $f_0'(-r)$ and $f_0'(r)$, respectively since $zf_0'(z)/f_0(z) = \psi(z)$. Thus, the extremes in the equation (3.2) are in terms of r alone and also lead to the sharp bounds. Consequently, we obtain the following distortion theorem of Ma-Minda [13] as a special case of Theorem 3.1 :

Let $\min_{|z|=r} |\psi(z)| = \psi(-r)$ *and* $\max_{|z|=r} |\psi(z)| = \psi(r)$. *If* $f \in \mathcal{S}^*(\psi)$ *and* $|z_0| = r < 1$. *Then*

$$f_0'(-r) = \psi(-r)\frac{f_0(-r)}{-r} \leq |f'(z_0)| \leq \frac{f_0(r)}{r}\psi(r) = f_0'(r).$$

Equality holds for some $z_0 \neq 0$ *if and only if* f *is a rotation of* f_0.

4. Bohr-Phenomenon for functions in $S_f(\psi)$

Note that "the Bohr radius of the class $\mathcal{X}$ is at least r_x", this holds for every result in this section. In general, Bohr radius is estimated for a specific class provided the sharp coefficients bounds of the functions in that class are known. For instance, consider the class of starlike univalent functions, where we have the sharp coefficient bounds: $|a_n| \leq n$. However, for most of the Ma-Minda subclasses, the better coefficients bounds are yet not known. Hence, we encounter the following problem, especially in context of Ma-minda classes, which we deal here to a certain extent:
Problem: *If coefficients bounds are not known, how one can find a good lower estimate for the Bohr radius of a given class?*

To readily understand the above problem, consider the class $\mathcal{S}^*(1 + ze^z)$, where the sharp coefficients bounds for functions in this class are unknown. In this situation, how one can find the Bohr radius for this class or is there any way out with the lower bounds all alone? Here below we state Theorem 4.1, where we find a solution for this problem. Note that the Bohr radius $3 - 2\sqrt{2} \approx 0.1713$ for the class $\mathcal{S}^*$ serves as a lower bound for the class $S_f(\psi)$ and is also a special case of Theorem 4.1.

Let $\mathbb{B}(0, r) := \{z \in \mathbb{C} : |z| < r\}$, $g(z) = \sum_{k=1}^{\infty} b_k z^k$, $\mathcal{S}^*(\psi)$ and $S_f(\psi)$ as defined in (1.1) and (1.3) respectively. For any $g \in S_f(\psi)$, we find the radius r_b so that $S_f(\psi)$ obey the following Bohr-phenomenon:

$$\sum_{k=1}^{\infty} |b_k| r^k \leq d(f(0), \partial\Omega) \quad \text{for} \quad |z| = r \leq r_b, \tag{4.1}$$

where $d(f(0), \partial\Omega)$ denotes the Euclidean distance between $f(0)$ and the boundary of $\Omega = f(\mathbb{D})$. Now we prove our main result:

Theorem 4.1. *Let* r_* *be the Koebe-radius for the class* $\mathcal{S}^*(\psi)$, $f_0(z)$ *be given by the equation* (3.1) *and* $g(z) = \sum_{k=1}^{\infty} b_k z^k \in S_f(\psi)$. *Assume* $f_0(z) =$

$z+\sum_{n=2}^{\infty} t_n z^n$ *and* $\hat{f}_0(r) = r+\sum_{n=2}^{\infty} |t_n| r^n$. *Then* $S_f(\psi)$ *satisfies the Bohr-phenomenon*

$$\sum_{k=1}^{\infty} |b_k| r^k \leq d(f(0), \partial\Omega), \quad \textit{for } |z| = r \leq r_b, \tag{4.2}$$

where $r_b = \min\{r_0, 1/3\}$, $\Omega = f(\mathbb{D})$ *and* r_0 *is the least positive root of the equation*

$$\hat{f}_0(r) = r_*.$$

The result is sharp when $r_b = r_0$ *and* $t_n > 0$.

Proof. Since $g \in S_f(\psi)$, we have $g \prec f$ for a fixed $f \in \mathcal{S}^*(\psi)$. By letting r tends to 1 in Lemma 3.1, we obtain the Koebe-radius $r_* = -f_0(-1)$. Therefore $\mathbb{B}(0, r_*) \subset f(\mathbb{D})$, which implies that

$$r_* \leq d(0, \partial\Omega) = |f(z)| \quad \text{for} \quad |z| = 1.$$

Also using the result [13, Theorem 1, p.161], we have

$$\frac{f(z)}{z} \prec \frac{f_0(z)}{z}. \tag{4.3}$$

Recall the result [2, Lemma 1, p.1090], which reads as: let f and g be analytic in $\mathbb{D}$ with $g \prec f$, where $f(z) = \sum_{n=0}^{\infty} a_n z^n$ and $g(z) = \sum_{k=0}^{\infty} b_k z^k$. Then $\sum_{k=0}^{\infty} |b_k| r^k \leq \sum_{n=0}^{\infty} |a_n| r^n$ for $|z| = r \leq 1/3$. Now using the result for $g \prec f$ and (4.3), we have

$$\sum_{k=1}^{\infty} |b_k| r^k \leq r + \sum_{n=2}^{\infty} |a_n| r^n \leq \hat{f}_0(r) \quad \text{for } |z| = r \leq 1/3.$$

Finally, to establish the inequality (4.2), it is enough to show $\hat{f}_0(r) \leq r_*$. But this holds whenever $r \leq r_0$, where r_0 is the least positive root of the equation $\hat{f}_0(r) = r_*$. The existence of the root r_0 is ensured by the relations

$$\hat{f}_0(1) \geq |f_0(1)| \geq r_* \quad \text{and} \quad \hat{f}_0(0) < r_*.$$

Thus, if $r_b = \min\{r_0, 1/3\}$ then $\sum_{k=1}^{\infty} |b_k| r^k \leq d(0, \partial\Omega)$ holds. The case of sharpness follows for the function f_0. □

Remark 4.1. Let us further assume that the coefficients B_n of ψ are positive. Then the function $f_0(z) = z+\sum_{n=2}^{\infty} t_n z^n$ defined by integral representation (3.1) can be written as

$$f_0(z) = z \exp\left(\sum_{n=1}^{\infty} \frac{B_n}{n} z^n\right),$$

which implies $f_0(r) = \hat{f}_0(r)$ for $|z| = r$.

From the proof of Theorem 4.1, we have the following:

Theorem 4.2. *Let r_* be the Koebe-radius for the class $\mathcal{S}^*(\psi)$, $f_0(z)$ be given by the equation (3.1) and $f(z) = z + \sum_{n=2}^{\infty} a_n z^n \in \mathcal{S}^*(\psi)$. Assume $f_0(z) = z + \sum_{n=2}^{\infty} t_n z^n$ and $\hat{f}_0(r) = r + \sum_{n=2}^{\infty} |t_n| r^n$. Then $\mathcal{S}^*(\psi)$ satisfies the Bohr-phenomenon*

$$r + \sum_{n=2}^{\infty} |a_n| r^n \leq d(f(0), \partial\Omega), \quad \text{for } |z| = r \leq r_b,$$

where $r_b = \min\{r_0, 1/3\}$, $\Omega = f(\mathbb{D})$ and r_0 is the least positive root of the equation

$$\hat{f}_0(r) = r_*.$$

The result is sharp when $r_b = r_0$ and $t_n > 0$.

If we choose $\psi(z) = (1 + Dz)/(1 + Ez)$, $-1 \leq E < D \leq 1$, then $\mathcal{S}^*(\psi)$ denotes the class of Janowski starlike functions and we have

$$r_* = \begin{cases} (1-E)^{\frac{D-E}{E}}, & E \neq 0; \\ e^{-D}, & E = 0. \end{cases} \tag{4.4}$$

and

$$f_0(z) = \begin{cases} z(1+Ez)^{\frac{D-E}{E}}, & E \neq 0; \\ z\exp(Dz), & E = 0. \end{cases} \tag{4.5}$$

Observe that if $E \neq 0$, the n-th ($n \geq 2$) coefficients of $f_0(z)$ is given by

$$t_n = \prod_{k=2}^{\infty} \frac{D - (k-1)E}{(k-1)!}. \tag{4.6}$$

Thus from Theorem 4.1, we have the following result:

Corollary 4.3. *Let $\psi(z) = (1 + Dz)/(1 + Ez)$, $-1 \leq E < D \leq 1$. Then $S_f(\psi)$ (and $\mathcal{S}^*(\psi)$) satisfies the Bohr-phenomenon (4.1) for $|z| = r \leq r_b$, where $r_b = \min\{r_0, 1/3\}$ and r_0 is the least positive root of the equation*

$$r + \sum_{n=2}^{\infty} |t_n| r^n - (1-E)^{\frac{D-E}{E}} = 0,$$

where t_n is as defined in (4.6).

Note that for the Janowski class, sharp coefficients bounds in general are not known. Now as an application of Corollary 4.3, we obtain the following result when $t_n > 0$:

Corollary 4.4. (Bohr-radius with Janowski class) *Let $\psi(z) = (1 + Dz)/(1 + Ez)$, $-1 \leq E < D \leq 1$.*

(*i*) *If $E = 0$ and $D \geq \frac{3}{4}\log 3$. Then $S_f(\psi)$ (and $\mathcal{S}^*(\psi)$) satisfies the Bohr-phenomenon (4.1) for $|z| = r \leq r_0$, where r_0 is the only real root of the equation*

$$1 - re^{D(1+r)} = 0. \tag{4.7}$$

(*ii*) *If* $E \neq 0$ *and further satisfies*

$$3(1-E)^{\frac{D-E}{E}} \leq (1+E/3)^{\frac{D-E}{E}}. \tag{4.8}$$

Then $S_f(\psi)$ *(and* $\mathcal{S}^*(\psi)$*) satisfies the Bohr-phenomenon* (4.1) *for* $|z| = r \leq r_0$*, where* r_0 *is the only real root of the equation*

$$(1-E)^{\frac{D-E}{E}} - r(1+Er)^{\frac{D-E}{E}} = 0. \tag{4.9}$$

The result is sharp for the function f_0 *defined in* (4.5).

Proof. (i): Since $E = 0$, we have $r_* = e^{-D}$. Moreover $\hat{f}_0(r) = f_0(r) = r\exp(Dr)$. Now we need to show

$$r\exp(Dr) \leq e^{-D} \tag{4.10}$$

or equivalently $T(r) := 1 - re^{D(1+r)} \geq 0$ holds for $r \leq r_0$. Which obviously holds for $\frac{3}{4}\log 3 \leq D \leq 1$. Since $d(f_0(0), \partial f_0(\mathbb{D})) = r_*$, therefore we see from inequality (4.10) that Bohr-radius is sharp for the function f_0 given by (4.5).

(ii): Proceeding as in case (i), it is sufficient to show the inequality

$$r(1+Er)^{\frac{D-E}{E}} \leq (1-E)^{\frac{D-E}{E}} \tag{4.11}$$

or equivalently $g(r) := (1-E)^{\frac{D-E}{E}} - r(1+Er)^{\frac{D-E}{E}} \geq 0$ holds for $r \leq r_0$. Which obviously follows whenever D and E satisfies (4.8). In view of the inequality (4.11), the sharp Bohr-radius is achieved for the function f_0 given by (4.5). □

Remark 4.2. (Bohr-radius with starlike functions of order α) Let $\psi(z) := (1+(1-2\alpha)z)/(1-z)$, where $0 \leq \alpha < 1$. We see $\mathcal{S}^*(\psi) := \mathcal{S}^*(\alpha)$ and for this class, we have

$$r_* = \frac{1}{2^{2(1-\alpha)}} \quad \text{and} \quad f_0(z) = \frac{z}{(1-z)^{2(1-\alpha)}}.$$

Observe that here $\hat{f}_0(r) = f_0(r)$. Now as an application of Corrollary 4.4, we obtain the result due to Bhowmik et al. [2], namely:

If $0 \leq \alpha \leq 1/2$*. Then* $S_f(\psi)$ *satisfies the Bohr-phenomenon* $\sum_{k=1}^{\infty} |b_k| r^k \leq d(f(0), \partial\Omega)$*, for* $|z| = r \leq r_b$*, where* $r_b = \min\{r_0, 1/3\} = r_0$ *and* r_0 *is the only real root of the equation* $(1-r)^{2(1-\alpha)}/r = 2^{2(1-\alpha)}$*. The result is sharp.*

Now form the above remark, in particular, we have:

Corollary 4.5. *If* $0 \leq \alpha \leq 1/2$*. Then the class* $\mathcal{S}^*((1+(1-2\alpha)z)/(1-z))$ *satisfies the Bohr-phenomenon* (4.1) *for* $|z| = r \leq r_0$*, where* r_0 *is the only real root of the equation*

$$(1-r)^{2(1-\alpha)}/r = 2^{2(1-\alpha)}.$$

The result is sharp. In particular, the Bohr radius for the class $\mathcal{S}^*$ *is* $3 - 2\sqrt{2} \approx 0.1713$.

If we choose $\psi(z) = \sqrt{1+z}$, then $\mathcal{S}^*(\psi) := \mathcal{SL}^*$, the class of lemniscate starlike functions and for this class we have:

$$f_0(z) = \frac{4z\exp(2\sqrt{1+z}-2)}{(1+\sqrt{1+z})^2} \quad \text{and} \quad r_* = -f_0(-1) \approx 0.541341. \tag{4.12}$$

Also in this case $\hat{f}_0(r) = f_0(r)$ and therefore, we obtain the following corollary:

Corollary 4.6. *The class $S_f(\psi)$ (and $\mathcal{SL}^*$), where $\psi(z) = \sqrt{1+z}$ satisfies the Bohr-phenomenon* (4.1) *for $|z| = r \leq 1/3$.*

If we consider $\psi(z) = 1 + ze^z$, then $\mathcal{S}^*(\psi) := \mathcal{S}^*_\wp$, the class of cardioid starlike functions introduced in [10] and for this class, we have:

$$f_0(z) = z\exp(e^z - 1) \quad \text{and} \quad r_* = -f_0(-1) \approx 0.531464. \tag{4.13}$$

Here we can also see that $\hat{f}_0(r) = f_0(r)$ and we obtain the following corollary:

Corollary 4.7. *The class $S_f(\psi)$ (and $\mathcal{S}^*_\wp$), where $\psi(z) = 1 + ze^z$ satisfies the Bohr-phenomenon* (4.1) *for $|z| = r \leq 1/3$.*

Ali et al. [1] also showed that the coefficient bound of a function in a class have a role in the estimation of the Bohr-radius. Observed that for each $f \in \mathcal{S}^*(\psi)$, the class $S_f(\psi)$ satisfies the Bohr-phenomenon for $r \leq \min(1/3, r_0)$, where r_0 is the least positive root of $\hat{f}_0(r) - r_* = 0$. Since $\mathcal{S}^*(\psi) \subset \bigcup_{f\in\mathcal{S}^*(\psi)} S_f(\psi)$, therefore the Bohr-radius for the class $\mathcal{S}^*(\psi)$ is $r \geq \min(1/3, r_0)$. In Corollary 4.7, we find $r_0 \approx 0.349681$ (an upper bound for Bohr radius), which is almost close to $1/3 \approx 0.33333$ and is the unique root of $f_0(r) - r_* = 0$. Moreover, the bound for the coefficients of the functions belonging to $\mathcal{S}^*_\wp$ and $\mathcal{SL}^*$ have been conjectured [10, 23] with the extremals given in (4.13) and (4.12) respectively. Thus by using Theorem 4.1 and the approach dealt in [1] (assuming that conjectures are true), we propose the following conjectures:

Conjecture 4.1. The Bohr-radius for the class $\mathcal{S}^*_\wp$ is $r_0 \approx 0.349681$ which is the unique root in $(0, 1)$ of the equation

$$re^{e^r} = e^{1/e}.$$

Conjecture 4.2. The Bohr-radius for the class $\mathcal{SL}^*$ is $r_0 \approx 0.439229$, which the unique root in $(0, 1)$ of the equation

$$e^2 r\exp(2\sqrt{1+r}-2) = (1+\sqrt{1+r})^2.$$

Conflict of interest

The authors declare that they have no conflict of interest.

References

[1] R. M. Ali, N. K. Jain and V. Ravichandran, Bohr radius for classes of analytic functions, Results Math. 74(4) (2019), Art. 179, 13 pp.

[2] B. Bhowmik and N. Das, Bohr phenomenon for subordinating families of certain univalent functions, J. Math. Anal. Appl. 462(2) (2018), 1087–1098.

[3] H. P. Boas and D. Khavinson, Bohr's power series theorem in several variables, Proc. Amer. Math. Soc. 125(10) (1997), 2975–2979.

[4] H. Bohr, A theorem concerning power series, Proc. London Math. Soc. 13(2) (1914), 1–5.

[5] T. Bulboacă and N. Tuneski, New criteria for starlikeness and strongly starlikeness, Mathematica **43(66)** (2001), no. 1, 11–22 (2003).

[6] N. E. Cho, V. Kumar, S. S. Kumar, V. Ravichandran, Radius problems for starlike functions associated with the sine function, Bull. Iranian Math. Soc. 45(1) (2019), 213–232.

[7] K. Gangania and S. S. Kumar, On Certain Generalizations of $\mathcal{S}^*(\psi)$, Comput. Methods Funct. Theory (2021), https://doi.org/10.1007/s40315-021-00386-5.

[8] P. Goel and S. S. Kumar, Certain Class of Starlike Functions Associated with Modified Sigmoid Function, Bull. Malays. Math. Sci. Soc. (2019) https://doi.org/10.1007/s40840-019-00784-y.

[9] F. H. Jackson, On q -functions and certain difference operator. Trans. R. Soc. Edinb. 46, 253–281 (1908)

[10] S. S. Kumar and G. Kamaljeet, A cardioid domain and starlike functions. Anal. Math. Phys. 11(2) (2021), 13 pp. https://doi.org/10.1007/s13324-021-00483-7.

[11] S. S. Kumar and G. Kamaljeet, On Geometrical Properties of Certain Analytic functions, Iran. J. Sci. Technol. Trans. A Sci. (2020), https://doi.org/10.1007/s40995-021-01116-1

[12] S. S. Kumar and G. Kamaljeet, Subordination and Radius Problems for Certain Starlike functions (communicated), arXiv:2007.07816v1.

[13] W. C. Ma and D. Minda: A unified treatment of some special classes of univalent functions, in *Proceedings of the Conference on Complex Analysis (Tianjin, 1992)*, 157–169, Conf. Proc. Lecture Notes Anal., I Int. Press, Cambridge, MA.

[14] R. Mendiratta, S. Nagpal, V. Ravichandran, On a subclass of strongly starlike functions associated with exponential function, Bull. Malays. Math. Sci. Soc. 38(1) (2015), 365–386.

[15] Y. A. Muhanna, R. M. Ali, Z. C. Ng and S. F. M. Hansi, Bohr radius for subordinating families of analytic functions and bounded harmonic mappings, J. Math. Anal. Appl. 420(1) (2014), 124–136.

[16] V. I. Paulsen, G. Popescu and D. Singh, On Bohr's inequality, Proc. London Math. Soc. 85(2) (2002), 493–512.

[17] K. Piejko and J. Sokół, On convolution and q-calculus, Bol. Soc. Mat. Mex. (3) **26** (2020), no. 2, 349–359.

[18] R. K. Raina and J. Sokół, Some properties related to a certain class of starlike functions, C. R. Math. Acad. Sci. Paris. 353(11), 973–978(2015)

[19] St. Ruscheweyh and T. Sheil-Small, Hadamard products of Schlicht functions and the Pólya-Schoenberg conjecture, Comment. Math. Helv. **48** (1973), 119–135.

[20] H. Silverman, E. M. Silvia and D. Telage, Convolution conditions for convexity starlikeness and spiral-likeness, Math. Z. **162** (1978), no. 2, 125–130.

[21] H. Silverman and E. M. Silvia, Subclasses of starlike functions subordinate to convex functions, Canad. J. Math. **37** (1985), no. 1, 48–61.

[22] H. Silverman, Radii problems for sections of convex functions, Proc. Amer. Math. Soc. **104** (1988), no. 4, 1191–1196.

[23] J. Sokół, Coefficient estimates in a class of strongly starlike functions, Kyungpook Math. J. 49(2) (2009), 349–353.

[24] G. Szegö, Zur Theorie der schlichten Abbildungen, Math. Ann. **100** (1928), no. 1, 188–211.
[25] A. W. Goodman and I. J. Schoenberg, On a theorem of Szegő on univalent convex maps of the unit circle. J. Analyse Math. 44 (1984/85), 200–204.

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India

Email address: gangania.m1991@gmail.com

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India

Email address: spkumar@dce.ac.in

Journal of Molecular Structure 1243 (2021) 130857

Contents lists available at ScienceDirect

Journal of Molecular Structure

journal homepage: www.elsevier.com/locate/molstr

Penta -hexa coordination behaviour of ABA-P(V) corrole

Ritika Kubba[a], Omprakash Yadav[b], Pinky Yadav[c], Natalia Fridman[d], Anil Kumar[a,*]

[a] Department of Applied Chemistry, Delhi Technological University, Delhi-42, India
[b] Department of Chemistry, ARSD College, University of Delhi, Delhi, India
[c] Department of Chemistry, Graduate School of Basic Sciences, Hiroshima University, Japan
[d] Schulich Faculty of Chemistry, Technion-Israel Institute of Technology, Haifa, Israel

ARTICLE INFO

Article history:
Received 11 May 2021
Revised 30 May 2021
Accepted 3 June 2021
Available online 9 June 2021

Keywords:
Corrole
Phosphorous
NMR
Penta-hexa

ABSTRACT

ABA (where A = nitrophenyl, and B = pentafluorophenyl group) corrole was isolated as hexacoordinate P(V) complexes that contained two axial hydroxyl groups. NMR spectroscopy (^{1}H, ^{19}F and ^{31}P) in $CDCl_3$ indicated that the hexacoordinated P(V) complexes were prone to axial-ligand dissociation to form pentacoordinated P(V) complexes. Although, in the presence of strong coordinating solvents, such as CH_3OH and CD_3OD, the P(V) corrole complexes were observed to exist in a hexacoordinated geometry in which the strong solvent molecules acted as the ligands at the axial position. Also, we report here the existence of the hexacoordinated P(V) complexes in non-coordinating solvents such as $CHCl_3$, CH_2Cl_2 by using fluoride and methyl substituted silyloxy groups acting as the ligands at the axial position. Moderately bulky nature and electron-donation effect of silyloxy groups proved to be of great aid to prevent the axial ligand dissociation of silyloxy substituted hexacoordinated P(V) complex and strong P-F bond in non-coordinating solvents. The structure of hexacoordinate P(V) complex containing two axial -OH groups and hexacoordinate P(V) complex containing two axial -OCH_3 groups were confirmed by X-ray crystallography.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

Corroles are analogues of porphyrin, when compared to porphyrins they have one carbon atom less in their outer fringe and one NH proton extra in the inner cavity [1]. Metal ions in higher oxidation states can be stabilized by corroles due to the presence of -3 charge in their deprotonated form, whereas porphyrins on the other hand have -2 charge in their deprotonated form and thus can stabilize metals in low oxidation states [2]. This makes corroles more interesting, when it comes to study of metallocorroles with metal ions in higher oxidation states. Transition metal complexes of corroles are very well known and studied in much detail but only few researches have been made on metallocorroles containing main group elements [3-7]. Gross and coworkers [8] have reviewed the corrole complexes having main group elements in their cavity and discussed their synthetic protocols, spectroscopic and electrochemical aspects along with their stability and chemical reactivity. High fluorescent behaviour of corrole with main group elements makes their study of great importance and of much attraction due to high recorded quantum yield for Ga (III) [9-11], Al (III) [9-11], Sn (IV) [12-14] and P(V) [15-19] corroles. A wide variety of applications including photocatalysis [20], dye-sensitized solar cells (DSSC) [21], photodynamic therapy (PDT) for cancers [22,23], photodynamic inactivation (PDI) of microorganisms [24], bio-cell-imaging [25] and triplet-triplet upconversion [26] have been explored so far. Phosphorus being smaller in size as compared to usual metal ions proved to be a perfect fit for the corrole cavity. In 1998 Paolesse and coworkers [15] first reported the phosphorus corrole complex followed by Kadish and co-workers in 2000[19] by reacting the free base corroles and phosphorus carriers, such as $POCl_3$ and PCl_3, under reflux atmosphere using pyridine as a solvent. In 2012, Ravikanth and co-workers reported a library of *meso*-triaryl P(V) corroles and discussed the existence of penta- or hexacoordinated P(V) corroles [27]. A severe issue of ligand dissociation was faced and overcome by modifying the ligand type or the coordinating solvents [28,29]. In 2016, Ravikanth and co-workers took the support of axial silyloxy groups to avert the axial ligand dissociation of P(V) corrole complexes to obtain the hexacoordinated complex for P(V) corroles in non-coordinating solvents. Recently, Gross and co-workers have discussed novel approach for P(V) corrole functionalization by means of acid catalysed axial ligand interchange [30]. Corroles reported so far are usually of two types namely A_3 type and ABA type [1]. Where A_3 type corresponds to corroles with same meso substituent at C_5, C_{10} and C_{15} positions on corrole and ABA type corroles are those where same meso sub-

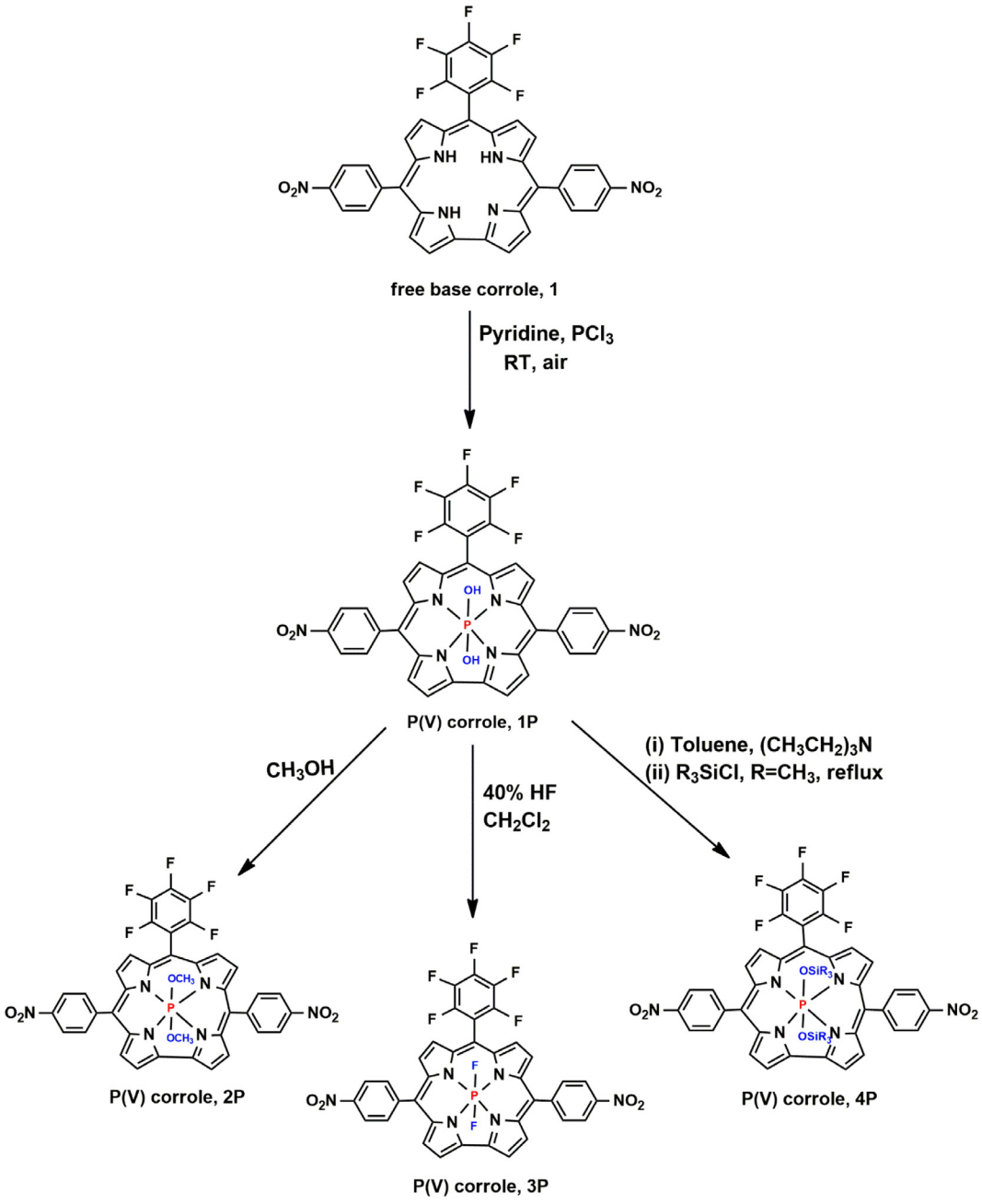
* Corresponding author.
E-mail address: anil_kumar@dce.ac.in (A. Kumar).

https://doi.org/10.1016/j.molstruc.2021.130857
0022-2860/© 2021 Elsevier B.V. All rights reserved.

stituent are present on the C_5 and C_{15} positions and different meso substituent is present on the C_{10} position on corrole. Axial ligand dissociation studies reported by Ravikanth et al. [28,29] were limited to P(V) corroles of A_3 type and to best of our knowledge no such report has been there for ABA type P(V) corroles. In this article, we have synthesized P(V) ABA corrole with electron withdrawing group at C_{10} meso position and studied coordination behaviour. This behaviour was well established with NMR and X-ray crystallography studies.

2. Results and Discussions

The insertion of phosphorus into the free base corrole namely 10-(pentafluorophenyl)-5, 15-bis(4-nitrophenyl) corrole (**1**) cavity has been carried out using $POCl_3$ or PCl_3 in the past. We have used a very facile and high yielding methodology introduced by Gross and coworkers [31] where PCl_3 has been used as the metal carrier and was added to the refluxing mixture of free base corrole and pyridine at room temperature under air (Scheme 1). The reaction progress was checked by taking the TLC. The reaction mixture changed its color from dark green to fluorescent green and then finally a reddish pink colored product was obtained with greenish tint. Due to the presence of pyridine and PCl_3 in the reaction mixture, we have observed little precipitation. In order to remove this precipitate, after successful addition of required amount of PCl_3 we have added water to the reaction mixture to dissolve the precipitate in water. Thereafter, standard work up was carried out using dichloromethane (reaction product dissolved in dichloromethane) and the excess solvent was evaporated under reduced pressure on rotary evaporator. The solid collected was subjected to column chromatography on silica gel to obtain the desired product, 10-(pentafluorophenyl)-5, 15-bis(4-nitrophenyl)



Scheme 1. Synthetic scheme of Phosphorus corrole, compound **1P, 2P, 3P** and **4P.**

phosphorus(V) corrole (**1P)** as reddish-pink coloured solids in 75% yield.

Titration of **1P** was carried out in strongly coordinating solvent *viz*. methanol, in excess of methanol, **1P** resulted into hexacoordinated compound namely 10-(pentafluorophenyl)-5, 15-bis(4-nitrophenyl) phosphorus(V) corrole, $(OCH_3)_2$ (**2P**) (Scheme 1). The exchange of axial –OH was also performed on stirring the solution of **1P** in dichloromethane and excess amount of HF for overnight. After overnight stirring excess HF was neutralized with water and 10-(pentafluorophenyl)-5, 15-bis(4-nitrophenyl) phosphorus(V) corrole, $(F)_2$ (**3P)** (Scheme 1) was obtained as pinkish crystalline compound on column chromatography. The compounds **1P** and **3P** were found to be soluble in common organic solvents such as CH_2Cl_2 and $CHCl_3$ and were characterized and confirmed by HR-MS, UV-visible, and detailed NMR spectroscopy. Further, **1P** and **2P** were also characterized by single crystal X-ray crystallography.

We realized that the hexacoordinated P(V) corrole can be obtained in the presence of coordinating solvent, like CH_3OH which eventually replaces the axial –OH groups. Whereas we were not able to obtain the hexacoordinated species in non-coordinating solvents, like $CHCl_3$ and CH_2Cl_2, reason being the vulnerability of the axial hydroxyl ligand to dissociate and result into penta and hexacoordinated species. We can suggest that, because of weak binding between axial hydroxyl groups (–OH) and the central phosphorus atom, a mixture containing both penta- and hexacoordinated P(V) corrole complexes have been found to exist.

Our interest towards obtaining hexacoordinated species in non-coordinating solvents, motivated us to synthesize 10-(pentafluorophenyl)-5, 15-bis(4-nitrophenyl) phosphorus(V) corrole, $(OSi(CH_3)_3)_2$ (**4P)** (Scheme 1). Due to the presence of a much bulkier group (siloxyl groups) at the axial position we successfully obtained the desired hexa-coordinated P(V)corrole (**4P**). Basically, the $p\pi$–$d\pi$ overlap between the silicon and oxygen atoms helps in strengthening the axial phosphorus-oxygen bonds and further prevents the ligand dissociation at the axial position.

The axial silyloxy substituted *meso*--triaryl P(V) corrole **4P** was prepared by reacting compound **1P** with chlorotrimethylsilane $[(CH_3)_3SiCl]$ in the presence of a base under reflux for 20-25 min (Scheme 1) in toluene as a solvent. To monitor the reaction progress TLC was checked frequently. Slight color change was witnessed from pink red to orange red as the reaction progressed. After the reaction was completed, the solvent was evaporated under reduced pressure on rotary evaporator and the collected solids were subjected to column chromatography on silica gel to obtain the required compound **4P** as orange-red solid in 70% yield.

Compounds **1P** and **3P** were further characterized by HRMS spectra (See Figure S1 and S2 in the Supporting Information), which showed molecular-ion peaks corresponding to hexacoordinated species and $[M-17]^+$ peaks corresponding to pentacoordinate species, which possibly have been formed as a result of the removal of one axial OH group in case of compound **1P**. The compound **3P** molecular ion peak appears at 772.1681 in APCI negative mode.

3. NMR characterization

All the synthesized compounds (**1P-4P**) were characterized by detailed NMR spectroscopy. ^1H NMR and ^{31}P NMR techniques were very crucial for understanding the structural composition of synthesized phosphorus complexes. ^1H NMR ($CDCl_3$) and ^{19}F spectra of compound **1P** (shown in Fig. 1) were analysed. ^1H NMR ($CDCl_3$) of compound **1P** showed eight sets of doublets (δ= 8.7 ppm to 9.5 ppm) due to β-pyrrolic protons, four sets of doublets (δ= 8.68, 8.9, 8.98, 9.29 ppm) are quite intense signals correspond to hexacoordinated while other four sets of doublets (δ= 8.79, 9.10, 9.15, 9.45 ppm) are with trivial intensity are due to pentacoordinated geometry. This suggested that the synthesized compound might have undergone axial-ligand dissociation, that is from hexacoordinated to possibly pentacoordinated geometry and that is the reason why we have observed four sets of doublets due to pentacoordinated geometry and other four due to hexacoordinated geometry. The other two doublets observed at δ= 8.63 ppm and 8.43 ppm are due to presence of ortho- and meta- proton present on the nitrophenyl moiety in the compound **1P**. To further support the discussion of penta and hexa coordination behaviour, we took the aid of ^{31}P NMR. As per literature, ^{31}P NMR($CDCl_3$) helps to analyse the coordination number of the central atom namely phosphorus in various phosphorus complexes [28]. If the compound shows a peak at ≈ -100 ppm, then phosphorus in the centre is pentacoordinated but if the peak is at ≈ -200 ppm, then central phosphorus is hexacoordinated. ^{31}P NMR spectra ($CDCl_3$) of compound **1P** (shown in Fig. 1 (b).) showed, an intense peak at ≈ -192 ppm and a much less intense peak at ≈ -102 ppm, which confirmed the presence of hexacoordinated and pentacoordinated phosphorus center, respectively. Further in ^{19}F NMR, we measure two sets of fluorine signals due to penta and hexa coordinated geometries (Fig. 1(c)). The intensity of the signals confirms the effect of electron-withdrawing nature of *p*-NO_2 group present on *para* position of *meso*-phenyl groups (which in turn are present on C_5 and C_{15} of the corrole complex) which try to prefer hexacoordinated geometry rather pentacoordinated geometry as observed earlier in the case of electron-releasing group [28]. Furthermore, the single crystal structure of **1P** shows the stability of hexacoordinated over pentacoordinated in solid state. ^{31}P NMR spectrum studies ($CDCl_3$) indicated that compounds **1P-3P** are much more stable in their hexacoordinated geometry than pentacoordinated and experience slight ligand dissociation at the axial position. Thus, the vulnerability of the hexacoordinated species to axial-ligand-dissociation was stipulated by ^1H and ^{31}P NMR studies of compound **1P** in $CDCl_3$.

Above results suggested that the axially bound hydroxyl groups can easily dissociate in the solution due to weak bonds with the central P(V) ion. At this point, we began to think about stabilizing this hexacoordinated species and functioned our approach towards modifying the coordinating solvent.

We performed the titration of compound **1P** with a more coordinating solvent, CH_3OH with an increasing amount of it; ^1H and ^{31}P NMR spectra were obtained for all the sets. During the titration, in ^{31}P NMR we witnessed various major and minor signals which corresponded possibly to different intermediate species formed in the process: $[1P\ (OH)]^+$, $[1P\ (OH)(OCH_3)]$, $[1P\ (OH)_2]$ and $[1P\ (OCH_3)_2]$. As we increased the amount of CH_3OH, the peak corresponding to hexacoordinated species (≈ -173 ppm) in the ^{31}P NMR (shown in Fig. 2) started getting more prominent, and eventually the signal at ≈ -100 ppm corresponding to pentacoordinate species disappeared.

On the other hand, ^1H NMR (shown in Fig. 3) spectra experienced decrease of eight sets of β-pyrrolic proton signals to four sets of signals. We observed gradual disappearance of four sets of doublets (δ= 8.79, 9.10, 9.15, 9.45 ppm) which corresponded to the pentacoordinated geometry, and at the same time we noticed increase in the intensity of the other four sets of doublets (δ= 8.68, 8.9, 8.98, 9.29 ppm), disappearance of signal at -4.72 and appearance of signal at -2.67 ppm corresponding to hexacoordinated geometry in the ^1HNMR spectrum (Fig. 3)

From the above results, we deduced that CH_3OH evidently resulted in the stabilization of hexacoordinate species by replacing the weakly bound axially placed OH ligands with more strongly bound OCH_3 ligands through well described mechanism as reported earlier [28].

In order to further explore the behaviour of phosphorus complex **1P** in strongly coordinating solvent- CD_3OD, a reaction of complex **1P** with deuterated methanol *viz.* CD_3OD was carried

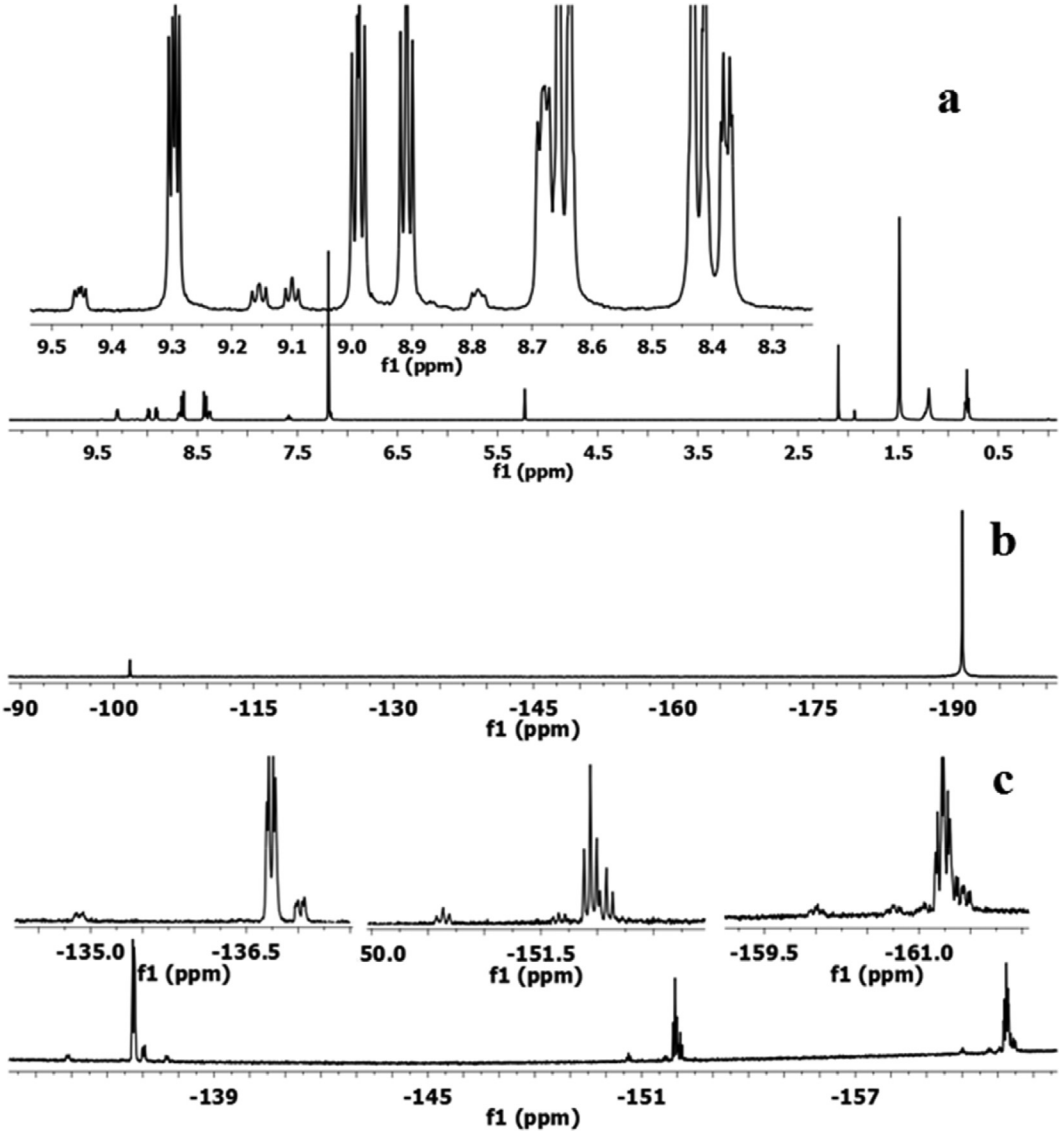


Fig. 1. (**a**) ^{1}H; (**b**) ^{31}P NMR and (**c**) ^{19}F NMR spectra of compound **1P** recorded in $CDCl_3$.

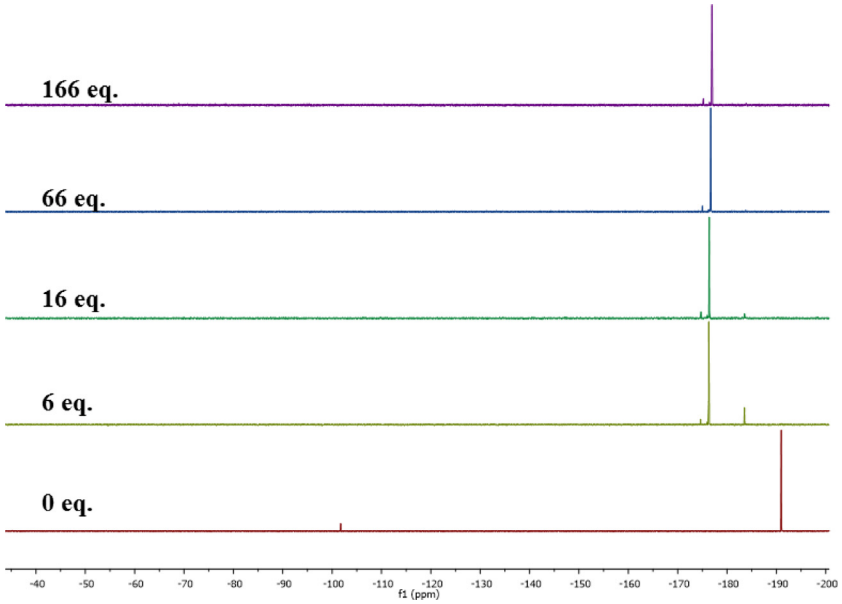


Fig. 2. ^{31}P NMR titration spectra for compound **1P** in $CDCl_3$ with addition of increasing amounts of CH_3OH in $CDCl_3$.

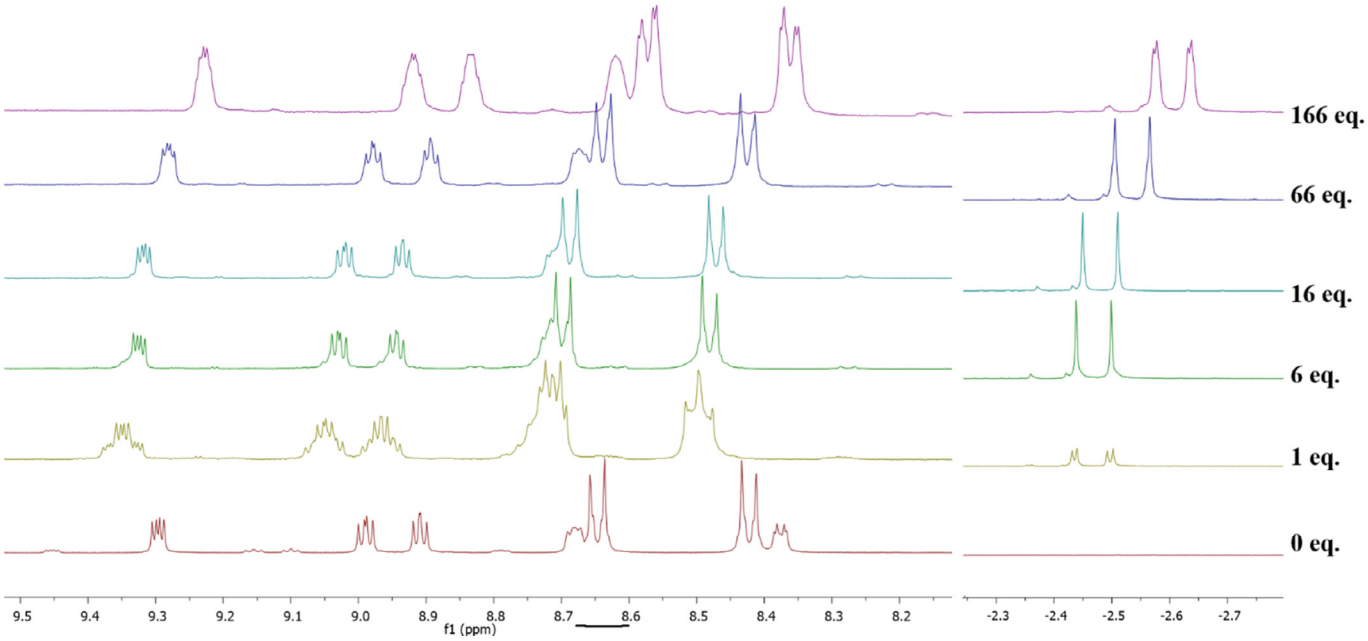


Fig. 3. ^{1}H NMR titration spectra for compound **1P** in $CDCl_3$ with addition of increasing amounts of CH_3OH in $CDCl_3$.

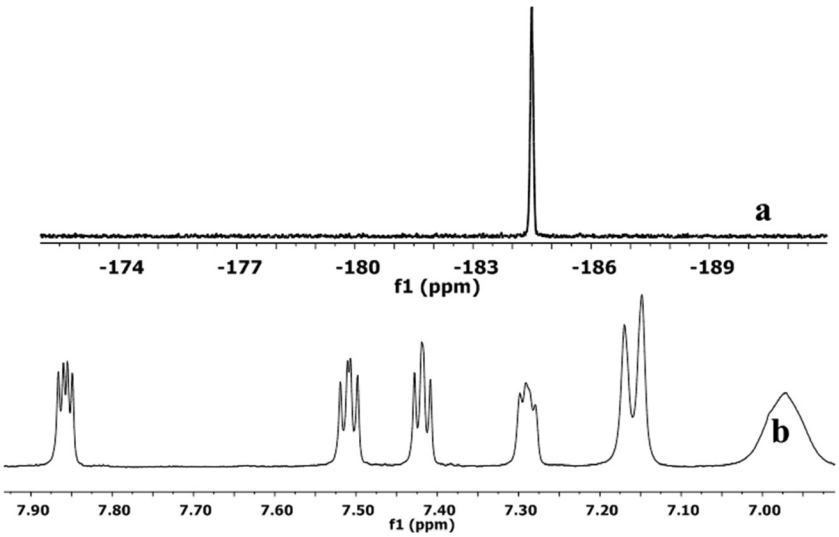


Fig. 4. (**a**) ^{31}P NMR and (**b**) ^{1}H NMR spectra of compound **1P** recorded in CD_3OD at room temp.

out. The existence of hexacoordinated phosphorus complex **1P** was confirmed using coordinating solvent: CD_3OD. ^{1}H NMR (in CD_3OD) showed four sets of doublets, each corresponds to two proton (δ= 7.28, 7.42, 7.52, 7.86 ppm) due to β-pyrrolic protons, two sets of doublets (δ= 7.15 ppm and 6.98 ppm) due to four ortho- and four meta- proton present on the nitrophenyl moiety. ^{31}P NMR (in CD_3OD) showed single signal at ≈ -184.5 ppm (Fig. 4) and further confirmation of the hexacoordinated geometry with OCH_3 as axial ligand is the appearance of doublet at δ = - 2.47 ppm, which confirm the substitution of the OH with OCH_3 during the titration of **1P** with methanol (Fig. 3).

To overcome another challenge of obtaining hexacoordinated phosphorus complex in non-coordinating solvents like $CDCl_3$, we synthesized the compound **4P.** This synthesized complex having a bulky axial silyloxy group prevented axial ligand dissociation of this phosphorus corrole complex. The existence of hexacoordinated phosphorus complex **4P** was confirmed by ^{31}P NMR (in $CDCl_3$) that showed a single signal at ≈ -200 ppm and shown in Fig. 5. To gain more insight into this phenomenon we synthesized the di-fluoro ligated phosphorous complex of **1P** and **3P.** ^{1}H NMR, ^{31}P NMR and ^{19}F NMR suggested the complete ligation of fluoro with phosphorous(V) and stable complex formation was achieved. The appearance of signals at δ= 8.67, 8.88, 9.03 ppm are due to β-pyrrolic protons in ^{1}H NMR, only one signal in ^{31}P NMR at ≈ -182 ppm and appearance of doublet in ^{19}F NMR at -36.65 ppm with 815.30 Hz coupling constant suggested the fully formation of hexacoordinated complex and shown in Fig. 6.

4. Absorption and fluorescence studies

The absorption spectra of synthesized free base corrole **1** and its corresponding phosphorous derivatives **1P and 3P** were recorded in DCM as solvent (Table 1) and shown in Fig. 7. Generally, free base triarylcorroles show three absorption bands in total, out of which two weak absorption bands are seen in the Q-band region and one strong and sharp absorption band is seen as a Soret band. Whereas P(V) triaryl corrole when compared with free base triarylcorroles exhibited different absorption characteris-

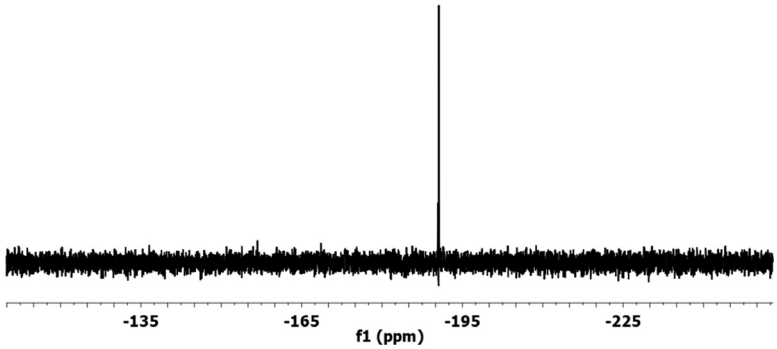


Fig. 5. ^{31}P NMR of **4P** in $CDCl_3$.

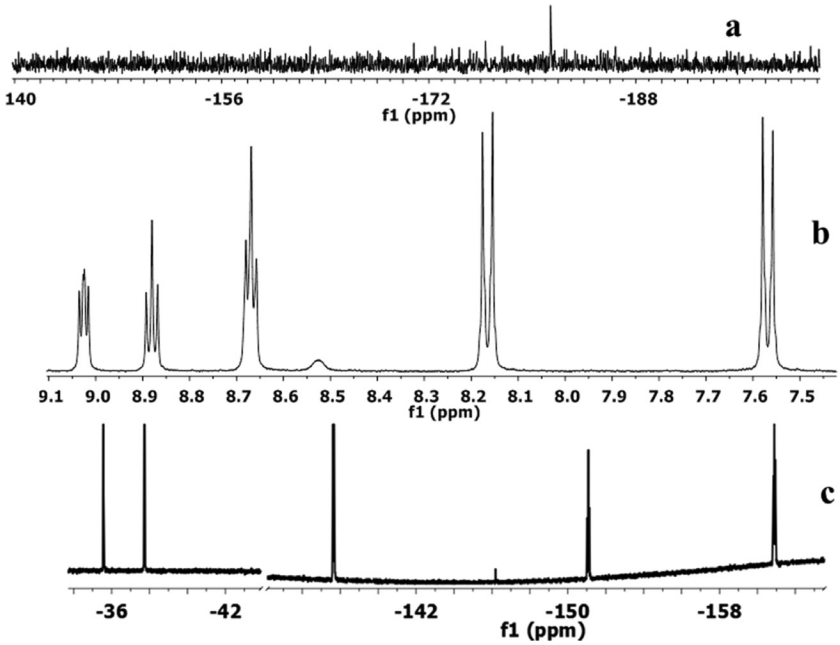


Fig. 6. a) ^{31}P NMR b) ^{1}H NMR and c) ^{19}F NMR of **3P** in C_6D_6.

Table 1
Absorption data of compound **1P, 3P** and **4P** recorded in dichloromethane and quantum yield of **1P, 3P** in toluene at room temperature.

Compound	Absorption data (in DCM)		Quantum Yield (%) in Toluene	Life time (ns)
	Soret band [nm] (log ε)	Q band [nm] (log ε)		
1P	425(5.28)	525(4.28) 565(4.36) 598(4.66)	23	3.4
3P	422(5.11)	520(4.03) 560(4.05) 590(4.33)	18	3.6
4P	413(4.42)	522(3.47) 571(3.54) 589(3.87)	–-	–-

tics. As shown in Fig. 7, compound **1P** exhibited strong Soret bands at ≈400 nm. However, unlike compound **1**, compound **1P** exhibited a much more intense band at about 600 nm with three Q-bands. Also, the absorption bands of compound **1P** was found to have undergone hypsochromic shift in Soret region compared to compound **1**. Hypsochromic shift was also observed with exchange of axial ligand with fluoride, **3P**.

Absorption spectra of compounds **1P** and **4P** were compared as shown in Fig. 8 (absorption data are reported in Table 1 of both the compounds). The absorption spectrum of compound **1P** depicts four absorption bands, out of which three Q-bands were seen at 525 nm, 565 nm, and 598 nm and one sharp Soret band was seen at 425 nm. On the other hand, the absorption spectrum of Compound **4P** witnessed similar absorption features as that of compound **1P**, with slight increase in their molar absorptivity and slim shifts in their peak maxima (Table 1).

Metallocorroles with main-group elements are strongly fluorescent and thus show higher values of quantum yields when com-

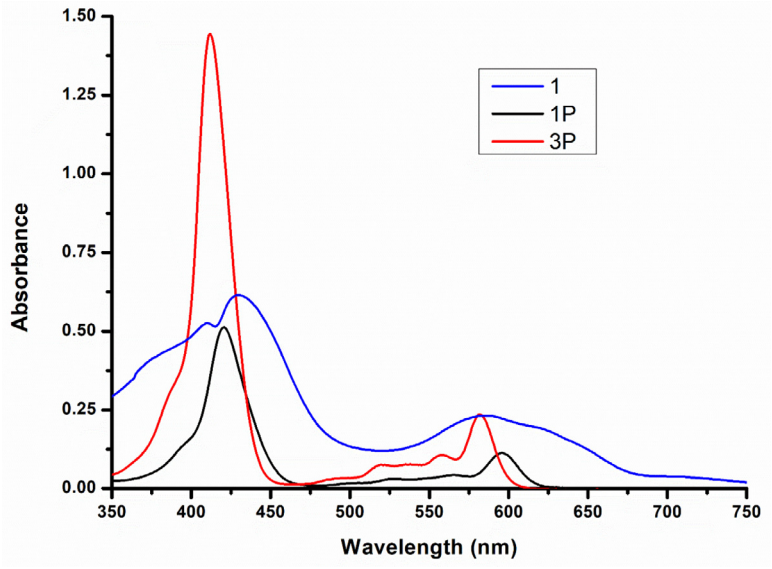


Fig. 7. Electronic spectra of compounds **1, 1P and 3P.**

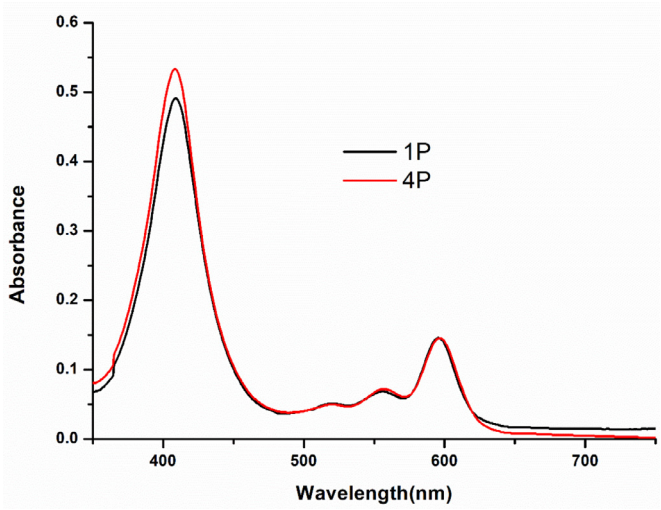


Fig. 8. Electronic spectra of compound **1P** and **4P** in DCM ($\approx 2 \times 10^{-5}$ M).

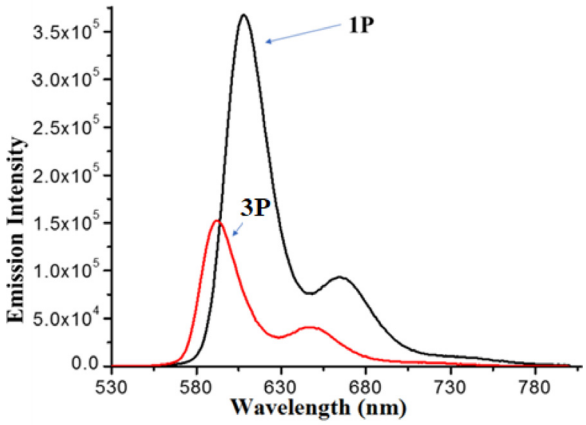


Fig. 9. Emission spectra of compounds **1P** and **3P** in toluene at room temperature.

pared with their porphyrin derivatives [9,32]. Photophysical properties of compounds **1P** and **3P** were studied (Table 1). Quantum yield of compound **1P** comes out to be lower than that of the previously reported phosphorus corroles [28,30]. Reason behind the lower value could be the presence of nitro group which quenches the fluorescence via spin orbit coupling mechanism which further enhances the rate of intersystem-crossing. Whereas, compound **3P** on the other hand exhibits higher quantum yield as compared to the reported phosphorus corroles [26]. A very strong fluorescence band around 600nm with shoulder bands were shown by compound **1P** (Fig. 9). Compound **3P** exhibited similar bands, but experienced a hypsochromic shift compared to **1P**.

5. Structural characterization

The formation of the compound **1P** and its methanol as axial ligand derivatives were confirmed by single-crystal X-ray crystallography. The ORTEPs view of compounds **1P** and **2P** are shown in Fig. 10. The **1P** was crystallized by slow diffusion of n-heptane into a dichloromethane solution of compound **1P** and **2P** in methanol solution resulting into red coloured crystals of **1P** and **2P. 1P** and **2P** crystallized in the monoclinic and triclinic space group $P2_1/c$ and P-1. Other crystallographic information is summarized in Table T1 (See the Supporting Information). Both the compounds **1P** and **2P** crystallized as a hexacoordinate complex with OH and OCH_3 groups as axial ligands. The deviation of the central phosphorus atom from the mean plane obtained by pyrrolic nitrogen's atoms was almost negligible. Usually in the pentacoordinate compound the phosphorus atom lies above the mean plane of corrole nitrogen's atoms and same was observed in metalloporphyrins and metallocorroles while in hexacoordinated complex prefers to be in plane [33-35]. Furthermore, the P-O bond lengths were found to be similar with other analogous P(V)–porphyrin and P(V)–corrole complexes [36]. The P-N bond distances of the **1P** and **2P** were inequivalent and the average P-N bond length (1.82 Å) was almost similar with metal complexes of meso-triarylcorroles reported [37,38]. Both the complexes **1P** and **2P** were obtained in planar geometries. Furthermore, in compounds **1P** and **2P** the meso-phenyl groups placed at B position make dihedral angle of 72.07° and 80.07° to the mean plane of the macrocycle. Interestingly, the axial hydroxyl ligand in the **1P** is opposite to each other likewise in another reported hydroxyl ligated hexacoordinated complex while in **2P** the OCH_3 group are not opposite to each other. This may be due to the steric relaxation of OCH_3 group

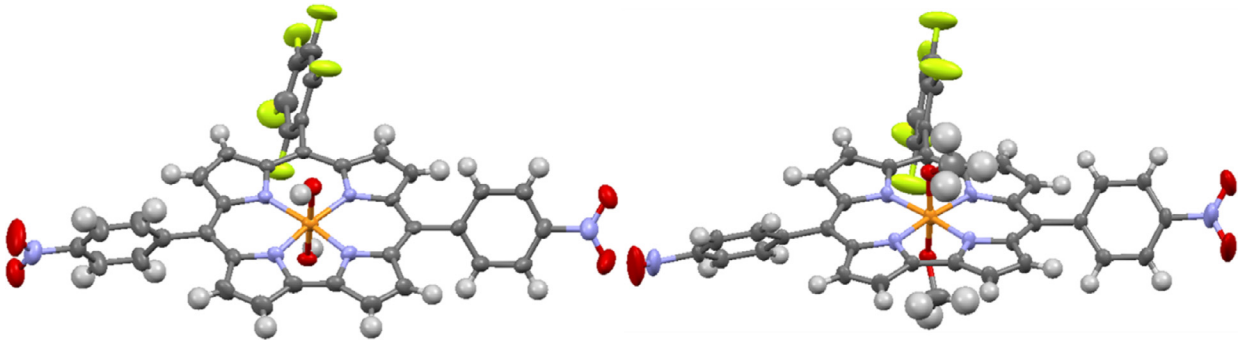


Fig. 10. ORTEP views of **1P** and **2P.** Thermal ellipsoids are shown at the 50% probability level.

with pentafluorophenyl group placed at C10 position. The detailed crystallographic information of **1P** and **2P** and their comparison with other metallocorroles are listed in Table T1.

6. Experimental Section

Chemicals: All chemicals were purchased from Sigma-Aldrich for synthesis of corroles and were used without further purification. Column chromatography of synthesised corroles was performed on silica gel (100–200 μm mesh) that was purchased from LOBA Chemie. Pyrrole was purified and distilled according to standard procedures. All solvents including dichloromethane, chloroform, toluene, n-heptane, hexane and methanol were purchased from E. Merck Ltd. Deuterated solvents ($CDCl_3$, CD_3OD and C_6D_6) were used as received from Sigma-Aldrich.

6.1. Instrumentation

6.1.1. Nuclear magnetic resonance spectroscopy

^{1}H NMR data was recorded on Jeol Model: JNM-EXCP 400 operating at 400 MHz spectrometer and Bruker Avance III 400 MHz. ^{19}F, and ^{31}P NMR spectra were recorded on a Bruker Avance III 400 spectrometer with automated tuning, operating at 400.4 MHz for ^{1}H, 376.7 MHz for ^{19}F, and 162.08 MHz for ^{31}P, respectively. Chemical shifts for hydrogens are reported relative to solvent peak at 7.26 ppm for $CDCl_3$.

6.2. UV–Visible spectroscopy

UV-1800 Shimadzu spectrophotometer was used for recording UV–Vis absorption spectra. Calculations of molar absorptivity coefficients of phosphorus corroles were done in dichloromethane.

6.3. High resolution mass spectrometry

High-resolution mass spectra were recorded on a Bruker Maxis Impact mass spectrometer, using APCI (atmospheric pressure charge ionization) in either positive or negative mode.

6.4. Fluorescence spectroscopy

Eemission spectra, emission quantum yield and life time measurements were recorded in toluene on a Fluorolog-3 spectrofluorometer (model FL3-11; HORIBA Jobin Yvon Inc., Edison, NJ) at room temperature. Tetraphenylpoprhyrin (TPP) was used as standard for quantum yield calculation [Φ_{tol} = 0.13] [39].

6.5. X-ray crystallography

Single crystals of **$C_{37}H_{18}F_5N_6O_6P$ (1-P)** were grown by slow evaporation of dichloromethane solution of **1-P** into heptane and single crystals of $C_{39}H_{22}F_5N_6O_6P$ (**2-P**) were grown by slow evaporation of methanolic solution of **2-P**. A suitable crystal was selected and mounted on a diffractometer (ApexDuo, Bruker-AXS, Mo Kα radiation). The crystal was kept at 200.15 K during data collection. Using Olex2 [40], the structure was solved with the olex2.solve [41] structure solution program using Charge Flipping and refined with the SHELXL [42] refinement package using Least Squares minimisation.

Crystal Data for $C_{37}H_{18}F_5N_6O_6P$ (1-P) (M =768.54 g/mol): monoclinic, space group $P2_1/c$ (no. 14), a = 10.964(2) Å, b = 33.810(6) Å, c = 10.598(2) Å, β = 99.986(5)°, V = 3869.2(12) Å [3], Z = 4, T = 200.15 K, μ(MoKα) = 0.429 mm^{-1}, *Dcalc* = 1.611 g/cm^3, 23533 reflections measured (2.41° ≤ 2Θ ≤ 49.146°), 5208 unique (R_{int} = 0.1092, R_{sigma} = 0.1491) which were used in all calculations. The final R_1 was 0.0893 (I > 2σ(I)) and wR_2 was 0.2648 (all data). **CCDC No**: **2082604.**

Crystal Data for $C_{39}H_{22}F_5N_6O_6P$ (2-P) (M =796.59 g/mol): triclinic, space group P-1 (no. 2), a = 8.4412(12) Å, b = 13.8734(18) Å, c = 15.360(2) Å, α = 72.291(4)°, β = 89.434(4)°, γ = 83.529(4)°, V = 1702.0(4) Å [3], Z = 2, T = 200.15 K, μ(MoKα) = 0.169 mm^{-1}, *Dcalc* = 1.554 g/cm^3, 15765 reflections measured (2.784° ≤ 2Θ ≤ 50.274°), 4045 unique (R_{int} = 0.1156, R_{sigma} = 0.1991) which were used in all calculations. The final R_1 was 0.0561 (I > 2σ(I)) and wR_2 was 0.1274 (all data). **CCDC No.: 2082605.**

Synthesis of compounds 1P-4P: Free-base corrole **1** was synthesized by using the literature procedures reported earlier [43] and characterization data were well matched (See Figure S3 and S4 in the Supporting Information) with previously reported same free base corrole synthesised in our lab [44].

For the synthesis of phosphorous corroles procedure is given below:

A solution of free base corrole (50 mg) in 10 mL pyridine was stirred vigorously for about 5 min. 300 μL of PCl_3 was added over 30 min (in 3 slots i.e. 100 μL every 10 min) to the reaction mixture with deep green color. Reddish-pink colored solution with a greenish red tint appeared after some time. Then to the reaction mixture 20 mL H_2O was added and was stirred strenuously for another 10min. The product was extracted via standard workup with CH_2Cl_2 and was washed thoroughly with H_2O. Green-red tint of the product became more prominent after extraction using CH_2Cl_2. UV-vis spectroscopy confirmed the completion of the reaction. The reaction mixture was evaporated under rotary evaporator pressure and subjected to column chromatography on silica. From a greenish reddish-pink-colored band the desired compound was collected by eluting with $CHCl_3$. The solvent was evaporated using a rotary evaporator to obtain a deep red-green solid. Given below are the analytical details for each compound.

10-(Pentafluorophenyl)-5, 15-bis(4-nitrophenyl) phosphorus(V) corrole, **1P**. The reddish-pink color solid was obtained in 75% yield. ^{1}H NMR (400 MHz, $CDCl_3$): δ 9.30 (dd, 4.5, 2.5 Hz, 2H), 8.99 (dd, 5.0, 3.62 Hz, 2H), 8.91 (dd, 4.3, 3.6 Hz, 2H), 8.67 (d, 5.1, 2H); ^{19}F

NMR (377 MHz, $CDCl_3$): −136.64 (dd, J = 7.12, 23.58, 2F; ortho-F), −151.08 (t, J = 21.97, 1F; para-F), −160.89 (dt, J = 5.65, 21.02, 2F; meta-F). ^{31}P NMR (400 MHz, $CDCl_3$): δ: -101.1 ppm, δ: -191.2 ppm; HRMS: m/z observed: 751.2333 [M-17] $^+$; calculated: 751.2500 [M-17] $^+$. UV–vis (CH_2Cl_2) λ_{max} (ε) [nm ($\times$ 10^6 cm^{-1} M^{-1})]: 425 (5.28), 525 (4.28), 565 (4.36), 598 (4.66).

10-(Pentafluorophenyl)-5, 15-bis(4-nitrophenyl) phosphorus(V) corrole, $(F)_2$ **3P**. ^{1}H NMR (400 MHz, C_6D_6): δ 9.03 (dd, 4.4, 3.3 Hz, 2H), 8.88 (t, 4.90 Hz, 2H), 8.67 (t, 4.5 Hz, 4H), 8.17 (d, 8.6 Hz, phenyl-H, 4H), 7.57 (d, 8.6 Hz, phenyl-H, 4H); ^{19}F NMR (377 MHz, $CDCl_3$): -37.16 (d, 815 Hz, 2F), −136.54 (dd, J = 7.12, 23.58, 2F; ortho-F), −151.58 (t, J = 21.97, 1F; para-F), −160.79 (dt, J = 5.65, 21.02, 2F; meta-F). ^{31}P NMR (400 MHz, $CDCl_3$): δ: -181.4 ppm; HRMS: m/z observed: 772.1681; calculated 772.5233; UV–vis (CH_2Cl_2) λ_{max} (ε) [nm ($\times$ 10^5 cm^{-1} M^{-1})]: 425 (5.11), 520 (4.03), 560 (4.05), 590 (4.33).

6.6. Synthesis of axial silyloxy P(V) corrole 4P

A solution of compound **1P** (25 mg) in dry toluene (31.32 mg) and dry triethylamine (34.40 mg) was stirred continuously for 5 minutes under N_2 atmosphere, to which chlorotrimethylsilane [$(CH_3)_3SiCl$, 14.77 mg] was added using syringe maintaining the N_2 atmosphere. The reaction mixture was stirred for another 20-25 minutes. The green-red tint color of the solution intensified during the reaction. The product was extracted via standard workup with CH_2Cl_2 and was washed thoroughly with H_2O. UV-vis spectroscopy confirmed the completion of the reaction. The reaction mixture was evaporated under rotary evaporator pressure and subjected to column chromatography on silica. From an orangish-red-colored band, the desired compound was collected by eluting with $CHCl_3$/Hexane (7:3). The solvent was evaporated using a rotary evaporator to obtain a deep orange-red solid in 70% yield. ^{31}P NMR ($CDCl_3$): δ: -190 ppm. UV–vis (CH_2Cl_2) λ_{max} (ε) [nm ($\times$ 10^4 cm^{-1} M^{-1})]: 413 (2.65), 522 (0.30), 571 (0.35), 589 (0.75).

7. Conclusion

We have shown here the penta and hexa coordination behaviour of (P(V)-ABA corrole. Characterization of synthesized corroles has been done by using various spectroscopic techniques. NMR spectroscopic studies have been carried in detail, which indicated that the P(V) *meso*--triarylcorrole (**1P**) conversely existed in hexacoordinate environments in coordinating solvents CH_3OH, CD_3OD. Whereas when in non-coordinating solvents (CH_2Cl_2, and $CHCl_3$), dissociated to their pentacoordinate complexes. Further, we also state that hexa coordinated species is the preferred geometry even in non-coordinating solvents. We also demonstrated the existence of hexa-coordinated geometry of P(V) corrole complexes even in non-coordinating solvents, such as CH_2Cl_2, by substituting axial hydroxyl groups of a central P(V) atom by methyl substituted silyloxy groups and fluoride atom. The electron withdrawing nature of *p*-NO_2 group prefer to hexacoordinated geometry as observed in the signal intensity in both ^{1}H and ^{31}P NMR.

CRediT author statement

Ritika Kubba: Writing- Original draft preparation, execution of the work, Methodology, **Anil Kumar:** Conceptualization, Methodology, Software and editing, **Om Prakash Yadav**: Reviewing, editing and conceptualization, **Pinky Yadav**: Methodology, Software and editing, Natalia Fridman: Crystal solving

Declaration of Competing Interest

There are no conflicts to declare.

Acknowledgements

We are grateful to Delhi Technological University (India) for infrastructural facilities. AK thanks to Professor Zeev Gross, Schulich Faculty of chemistry, Technion-Israel Institute of Technology, Haifa for using instrumental facilities.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.molstruc.2021.130857.

References

[1] I. Aviv, Z. Gross, Chem. Commun. 20 (2007) 1987.
[2] M. Saikawa, T. Noda, R. Matsuoka, T. Nakamura, T. Nabeshima, European Journal of Inorganic Chemistry 6 (2019) 766.
[3] R. Paolesse, S. Licoccia, M. Fanciullo, T. Boschi, Inorganica Chimica Acta 203 (1993) 107.
[4] L.M. Reith, M. Stiftinger, U. Monkowius, G. Knör, W. Schoefberger, Inorg. Chem. 50 (2011) 6788.
[5] A.M. Albrett, J. Conradie, P.D.W. Boyd, G.R. Clark, A. Ghosh, P.J. Bro- thers, J. Am. Chem. Soc. 130 (2008) 2888.
[6] K.M. Kadish, C. Erben, Z. Ou, V.A. Adamian, S. Will, E. Vogel, Inorg. Chem. 39 (2000) 3312.
[7] K.M. Kadish, S. Will, V.A. Adamian, B. Walther, C. Erben, Z. Ou, N. Guo, E. Vogel, Inorg. Chem. 37 (1998) 4573.
[8] I. Aviv, Z. Gross, Coord. Chem. Rev. 255 (2011) 717.
[9] M. Tasior, D.T. Gryko, J. Shen, K.M. Kadish, T. Becherer, H. Langhals, B. Ventura, L. Flamigni, Journal of Physical Chemistry C 112 (2008) 19699.
[10] D. Kowalska, X. Liu, U. Tripathy, A. Mahammed, Z. Gross, S. Hitayama, R.P. Steer, Inorg. Chem. 48 (2009) 2670.
[11] A. Mahammed, Z. Gross, J. Inorg. Biochem. 88 (2002) 305.
[12] L. Yun, H. Vazquez-Lima, H. Fang, Z. Yao, G. Geisberger, C. Dietl, A. Ghosh, P.J. Brothers, X. Fu, Inorg. Chem. 53 (2014) 7047.
[13] W. Sinha, S. Kar, Organometallics 33 (2014) 6550.
[14] W. Sinha, M. Kumar, A. Garai, C.S. Purohit, T. Som, S. Kar, Dalton Trans 43 (2014) 12564.
[15] R. Paolesse, T. Boschi, S. Licoccia, R.G. Khouryb, K.M. Smith, Chem. Commun. (1998) 1119.
[16] L. Simkhovich, A. Mahammed, I. Goldberg, Z. Gross, Chem. – Eur. J. 7 (2001) 1041.
[17] I. Saltsman, I. Goldberg, Z. Gross, Org. Lett. 17 (2015) 3214.
[18] J.P. Fox, D.P. Goldberg, Inorg. Chem. 42 (2003) 8181.
[19] K.M. Kadish, Z. Ou, V.A. Adamian, R. Guilard, C.P. Gros, C. Erben, S. Will, E. Vogel, Inorg. Chem. 39 (2000) 5675.
[20] A. Mahammed, Z. Gross, Angew. Chem. Int. Edit. 54 (2015) 12370.
[21] B.J. Brennan, Y.C. Lam, P.M. Kim, X. Zhang, G.W. Brudvig, ACS Appl. Mater. Interfaces. 7 (2015) 16124.
[22] R.D. Teo, J.Y. Hwang, J. Termini, Z. Gross, H.B. Gray, Chem. Rev. 117 (2017) 2711.
[23] Z. Zhang, H.J. Yu, H. Huang, H.H. Wang, S. Wu, H.Y. Liu, H.T. Zhang, Eur. J. Med. Chem. 163 (2019) 779.
[24] G. Light, Journal of Fungi 1 (2015) 138.
[25] X. Liang, J. Mack, L.M. Zheng, Z. Shen, N. Kobayashi, Inorg. Chem. 53 (2014) 2797.
[26] A. Mahammed, K. Chen, J. Vestfrid, J. Zhao, Z. Gross, Chemical Science 29 (2019) 7091.
[27] A. Ghosh, W.-Z. Lee, M. Ravikanth, Eur. J. Inorg. Chem. 26 (2012) 4231.
[28] A. Ghosh, M. Ravikanth, Chem. – Eur. J. 18 (2012) 6386.
[29] T. Chatterjee, W.-Z. Lee, M. Ravikanth, Dalton Transactions 18 (2016) 7815.
[30] Q.C. Chen, Z.Y. Xiao, S. Fite, A. Mizrahi, N. Fridman, X. Zhan, O. Keisar, Y. Cohen, Z. Gross, Chemistry - A European Journal 48 (2019) 11383.
[31] J. Vestfrid, R. Kothari, A. Kostenko, I. Goldberg, B. Tumanskii, Z. Gross, Inorganic Chemistry 12 (2016) 6061.
[32] a) A. Mahammed, Z. Gross, J. Inorg. Biochem. 88 (2002) 305; b) P.J. Brothers, Chem. Commun. (2008) 2090.
[33] a) A. Ghosh, T. Wondimagegn, A.B.J. Parusel, J. Am. Chem. Soc. 122 (2000) 5100; b) A. Ghosh, E. Steene, J. Biol. Inorg. Chem. 6 (2001) 740; c) I.H. Wasbotten, T. Wondi- magegn, A. Ghosh, J. Am. Chem. Soc. 124 (2002) 8104; d) Z. Gross, L. Simkhovich, N. Galili, Chem. Commun. (1999) 599; e) A.E. Meier–Callahan, A.J. Di Bi- lio, L. Simkhovich, A. Mahammed, I. Goldberg, H.B. Gray, Z. Gross, Inorg. Chem. 40 (2001) 6788; f) L. Simkhovich, Z. Gross, Tetrahedron Lett 42 (2001) 8089.
[34] a) R. Guilard, A. Zrineh, A. Tabard, A. Endo, B.C. Han, C. Le- comte, M. Souhassou, A. Habbou, M. Ferhat, K.M. Kadish, Inorg. Chem. 29 (1990) 4476; b) A.L. Balch, L. Latos-Graz˙yn´ski, B.C. Nom, S.L. Phillips, Inorg. Chem. 32 (1993) 1124; c) M.P. Schopfer, B. Mondal, D.H. Lee, A.A.N. Sarjeant, K.D. Karlin, J. Am. Chem. Soc. 131 (2009) 11304; d) R. Timkovich, A. Tulinsky, J. Am. Chem. Soc. 91 (1969) 16; e) V. Mckee, C.A. Rodley, Inorg. Chim. Acta. 151 (1988) 233.
[35] a) L. Simkhovich, N. Galili, I. Saltsman, I. Goldberg, Z. Gross, Inorg. Chem. 39 (2000) 2704; b) S. Nardis, R. Paolesse, S. Licoccia, F.R. Fronczek, M.G.H. Vicente, T.K. Shokhireva, S. Cai, F.A. Walker, Inorg. Chem. 44 (2005) 7030; c) L. Wagnert, A. Berg, E. Sta- vitski, T. Berthold, G. Kothe, I. Goldberg, A. Mahammed,

L. Sim- khovich, Z. Gross, H. Levanon, Appl. Magn. Reson. 30 (2006) 591; d) L. Simkhovich, P. Iyer, I. Goldberg, Z. Gross, Chem. Eur. J. 8 (2002) 2595; e) S. Nardis, F. Mandoj, R. Paolesse, F.R. Fronczek, K.M. Smith, L. Prodi, M. Montalti, G. Battistini, Eur. J. Inorg. Chem. (2007) 2345.

[36] a) C.A. Marrese, C.J. Carrano, Inorg. Chem. 22 (1983) 1858; b) C.J. Carrano, M. Tsutsui, J. Coord. Chem. 7 (1977) 79; c) P. Sayer, M. Gouterman, C.R. Connell, J. Am. Chem. Soc. 99 (1977) 1082.

[37] a) Y. Murakami, Y. Matsuda, S. Yamada, J. Chem. Soc. Dalton Trans. (1981) 855; b) Y. Matsuda, S. Yamada, K. Murakami, Inorg. Chem. 31 (1992) 2305; c) E. Vogel, S. Will, A.S. Tilling, L. Neumann, J. Lex, E. Bill, A.X. Trautwein, K. Wieghardt, Angew. Chem. 106 (1994) 771 Angew. Chem. Int. Ed. Engl. 33(1994) 731; d) S. Will, J. Lex, E. Vogel, V.A. Adamian, E.V. Caemelbecke, K.M. Kadish, Inorg. Chem. 35 (1996) 5577; e) E.V. Caemelbecke, S. Will, M. Autret, V.A. Adamian, J. Lex, J.P. Gisselbrecht, M. Gross, E. Vogel, K.M. Kadish, Inorg. Chem. 35 (1996) 184; f) M. Autret, S. Will, E.V. Caemelbecke, J. Lex, J.P. Gisselbrecht, M. Gross, E. Vogel, K.M. Kadish, J. Am. Chem. Soc. 116 (1994) 9141; g) J. Bendix, G. Golubkov, H.B. Gray, Z. Gross, Chem. Commun. (2000) 1957; h) S. Cai, F.A. Walker, S. Licoc- cia, Inorg. Chem. 39 (2000) 3466; i) K.E. Thomas, I.H. Wasbotten, A. Ghosh, Inorg. Chem. 47 (2008) 10469; j) B.O. Roos, V. Veryazov, J. Conradie, P.R. Taylor, Ghosh A, J. Phys. Chem. B 112 (2008) 14099.

[38] a) R. Paolesse, S. Licoccia, T. Boschi, Inorg. Chim. Acta 178 (1990) 9; b) J.W. Buchler, K.L. Lay, Inorg. Nucl. Chem. Lett. 10 (1974) 297; c) T. Barbour, W.J. Belcher, P.J. Brothers, C.E.F. Rickard, D.C. Ware, Inorg. Chem. 31 (1992) 746; d) L.M. Reith, M. Stiftinger, U. Monkowius, G. Knçr, W. Schoefberger, Inorg. Chem. 50 (2011) 6788; e) S. Licoccia, R. Paolesse, E. Tassoni, F. Poliziob, T. Boschi, Chem. Soc. Dalton Trans. (1995) 3617; f) A.M. Albrett, J. Conradie, P.D.W. Boyd, G.R. Clark, A. Ghosh, P.J. Brothers, J. Am. Chem. Soc. 130 (2008) 2888; g) D. Kowalska, X. Liu, U. Tripathy, A. Mahammed, Z. Gross, S. Hirayama, R.P. Steer, Inorg. Chem. 48 (2009) 2670; h) K.M. Kadish, C. Erben, Z. Ou, V.A. Adamian, S. Will, E. Vogel, Inorg. Chem. 39 (2000) 3312.

[39] D. Dolphin, The Porphyrins, III, Academic Press, New York, 1978.

[40] O.V. Dolomanov, L.J. Bourhis, R.J Gildea, J.A.K Howard, H. Puschmann, J. Appl. Cryst. 42 (2009) 339.

[41] L.J. Bourhis, O.V. Dolomanov, R.J Gildea, J.A.K Howard, H. Puschmann, Acta Cryst A71 (2015) 59.

[42] G.M. Sheldrick, Acta Cryst C71 (2015) 3.

[43] B. Koszarna, D.T. Gryko, J. Org. Chem. 71 (2006) 3707.

[44] O. Yadav, A. Varshney, A. Kumar, R.K. Ratnesh, M.S. Mehata, Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy 202 (2018) 207.



EasyChair Preprint

№ 5846

Performance Analysis of PV Arrays under Different Shading Conditions

Abhishek Yadav, Aaradhya Mittal, Nadeem Hussain Wazir, Vinod Kumar Yadav and Santosh Ghosh

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

June 17, 2021

Performance Analysis of PV Arrays under Different Shading Conditions

Abhishek Yadav[1], Aaradhya Mittal[1], Nadeem Hussain Wazir[1], Vinod Kumar Yadav[1], and Santosh Ghosh[2]

[1] Department of Electrical Engineering, Delhi Technological University, New Delhi, India-110042
[2] Corporate R&D Department, Kirloskar Brothers Limited, Pune, India-411045
vinodkumar@dtu.ac.in

Abstract. In order to fight climate change, investors, policy-makers, and power producers around the world are focusing on decreasing the share of generation from fossil fuels and increasing the share of renewable energy generation, especially through harnessing abundantly available solar energy. However, the power output and reliability of the photovoltaic (PV) plants are affected due to partial shading to a great extent. This paper investigates the effect of increasing the interconnection redundancy in the PV array in mitigative the power loss under different partial shading conditions in MATLAB® platform. The analysis revealed that the configuration with no interconnection redundancy (i.e., a series-parallel array) and with fill redundancy (i.e., a total cross-tied array) performs equally well under uniform irradiance and under self-shading created by low altitude sun, e.g., during sunrise and sunset, when the distance between adjacent rows is inadequate. In case of all other shading scenarios, such as shading created by nearby buildings, trees, poles, and non-uniform dust accumulation, an increase in the interconnection redundancy increases the performance ratio of the PV plant.

Keywords: Modeling and Simulation, Partial shading, Photovoltaics, PV array topology.

1 Introduction

The economic development of any country has a direct correlation with the quality and quantity of energy supply [1], but at the same time, the choice of energy resource plays a very crucial role in ensuring environmental sustainability [2]-[4]. Presently due to the global environmental crisis, globally there is a conscious effort to reduce the share of electricity generation from coal and other polluting fossil fuel and increase the use of solar energy, which is the most abundantly available clean energy resource [5] - [10]. However, solar energy has several inherent challenges [11], which impact its performance and as well as its reliability [12]; partial shading is one of those [13]. But most of the past studies revolve around the issue of enhancement of power output using different photo-responsive materials [14], [15], and the issues of

decline in output power and reliability due to partial shading conditions (PSC) have not been paid equal attention [16], [17]. Few methods have been developed to reduce the mismatch losses under PSC, such as multilevel converters [18], parallel connection of module [19], and micro converters [20], but a common downside among all these methods is the requirement of a large number of equipment, the increase in cost thereof, and the complexity of the algorithms [6],[7].

In this paper, a comparative study is presented on the performance of TCT configuration and conventional Series Parallel under different partial shading conditions. To investigate the performance of these two array configurations under different real-world scenarios of partial shading, in the present work, six shading patterns are simulated in MATLAB on a 3×3 PV array.

After this brief introduction, the rest of the paper is organized as follows. The modeling and simulation of PV cells are discussed in Section 2. In which the simulated shading modes and irradiance patterns are also delineated. Section 3 presents the results and discussion, followed by Section 4, which summarizes the main conclusions of the present work.

2 Modelling and Simulation

2.1 Modeling of PV Cell

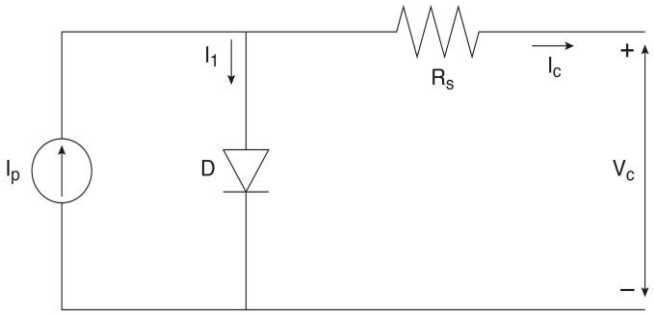


Fig. 1. Single diode equivalent circuit model of a PV cell.

An equivalent circuit diagram of PV cells is shown in Fig. 1, and the output current of the cell, *i.e.*, I_c may be quantified by Eq. (1) [13]. Here, I_p symbolizes photocurrent (i.e., current source), D denotes diode, Z denotes diode ideality factor, K symbolizes Boltzmann's constant, e is electron charge, while T and I_{sat} are temperature in °C, and saturation current, respectively.

$$I_c = I_p \ - I_{sat} \times \left(\frac{e(V_c + I_c R_s)}{zkT} - 1\right) \tag{1}$$

T_{op} denotes the operating temperature of a solar cell that depends on the solar irradiation level S_1. This is shown using Eq. (2) [13].

$$C_{Tg} = \ T \times (T_h - T_b) + 1 \quad \& \quad C_{Tg} = \frac{T \times (T_h - T_b)}{S_1} + 1 \tag{2}$$

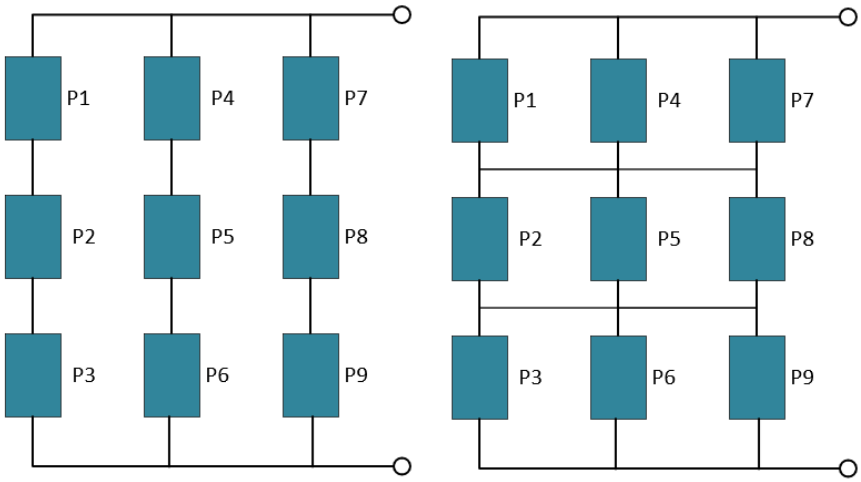
The effect of S_{ir}, the irradiation level on the voltage and photocurrent can be evaluated with the assistance of correction factors C_{Ss} and C_{So} as Eq. (3) [13].

$$C_{Ss} = T_S \times (S_{ir} - S_1) + 1, \quad C_{So} = \left(\frac{1}{S_1}\right) \times (S_{ir} - S_1) + 1 \tag{3}$$

S_1 and S_{ir} are the reference and insolation level, respectively.

2.2 The PV array and Shading modes

For investigating the relative performance of the series-parallel (SP) and total cross-ties (TCT) arrays, in the present work, a 3×3 matrix of PV modules is modeled in the MATLAB platform (see Fig. 2). In a real-world scenario, the PV arrays may be subjected to different shading conditions due to various environmental reasons, such as moving clouds, non-uniform dust accumulation over the panels, bird droppings, shading from nearby buildings, poles, trees, and other structures. To simulate these probable real-world scenarios, six different Cases are considered in the present study (see Table 1). The relative performances of both SP and TCT configurations are studied under these shading cases, and the result of the same is presented in Section 3.



(a) **(b)**

Fig. 2. PV configurations (a) Series-parallel and (b) total cross-tied.

3 Results and discussion

The P-V characteristics obtained under different shading conditions (see Table 1) are presented in Figs. 5 - 10. The exhibited P-V characteristics compare the results of the TCT circuit and the conventional series-parallel configuration. Tables 2 and 3 present

the performance parameters (*Pmax*, *Vmpp*, *Impp*) under different shading scenarios for the simulated circuits.

Table 1. Considered patterns and intensities over different modules of the array

	Irradiance, W/m2								
	P1	P2	P3	P4	P5	P6	P7	P8	P9
Case-1	1000	1000	1000	1000	1000	1000	1000	1000	1000
Case-2	400	600	200	1000	1000	1000	1000	1000	1000
Case-3	400	600	200	300	300	500	1000	1000	1000
Case-4	800	600	200	800	800	1000	200	1000	1000
Case-5	600	1000	1000	300	1000	1000	800	1000	1000
Case-6	500	300	1000	400	400	1000	300	500	1000

For Case-1, i.e., without shading scenarios (see Fig.3), the PV curves superimpose each other, and no difference could be found between SP and TCT arrays, which implies that both SP and TCT produce the same power under uniform irradiance condition. Cases 2 and 3, i.e., for vertical shading scenario, TCT shows better responses than SP and produces more power output (see Figs. 4 and 5). The multiple maxima may be observed in both of these cases due to partial shading conditions.

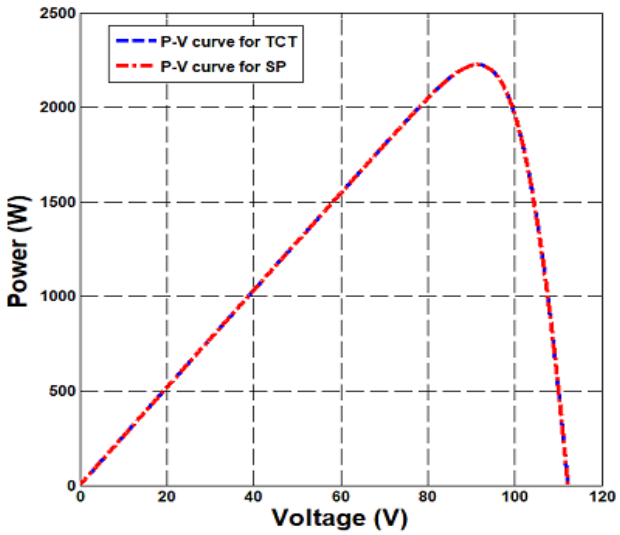


Fig. 3. *P-V* characteristics for Case 1.

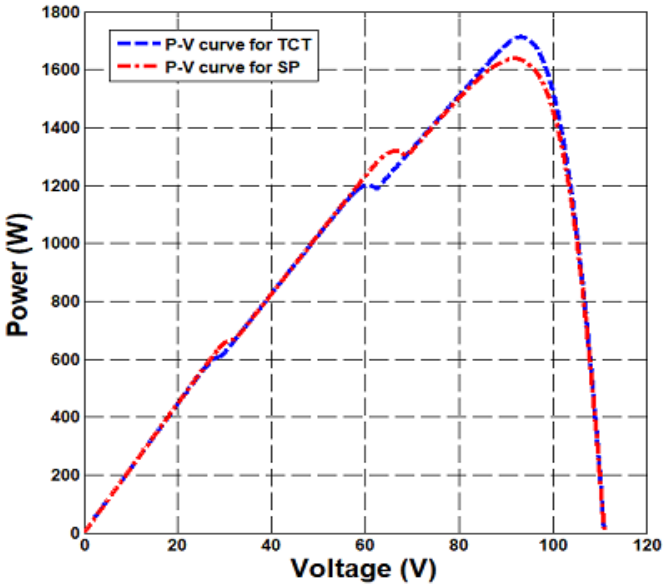


Fig. 4. *P-V* characteristics for Case 2.

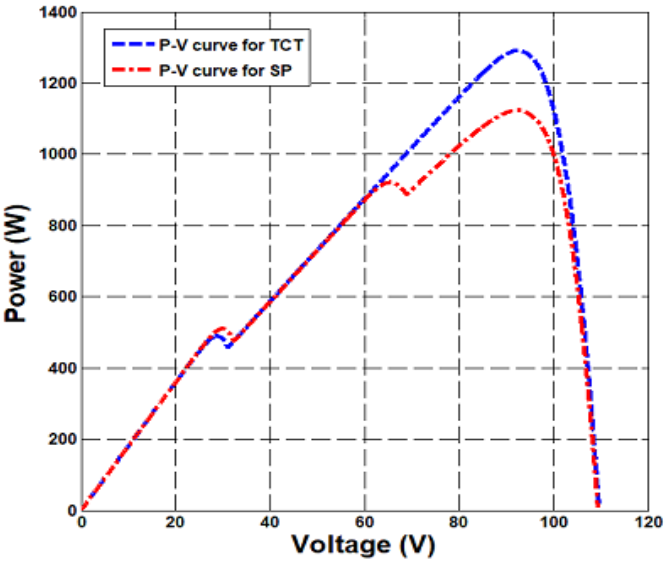


Fig. 5. *P-V* characteristics for Case 3.

Fig. 6 elucidate P-V characteristics under diagonal shading, i.e., under shading scenario represented by Case 4. In this case, the difference between the global peak power between the SP array and the TCT array increases to a much greater value (17.24 %) compared to the previous cases. Fig.7 represents the P-V characteristic for

Case 5, which is a horizontal shading scenario, and the results are of both the arrays are quite similar, and no difference could be found in the global peak points produced by SP and TCT arrays. Fig.8 represents P-V characteristics under Case 6, in which the upper two rows are under shading, making the column partially shaded. In this case, the TCT configuration produces relatively greater power output than that produced by SP array as well, but the absolute value of peak power is observed to be least in this case compared to other cases for both the configurations.

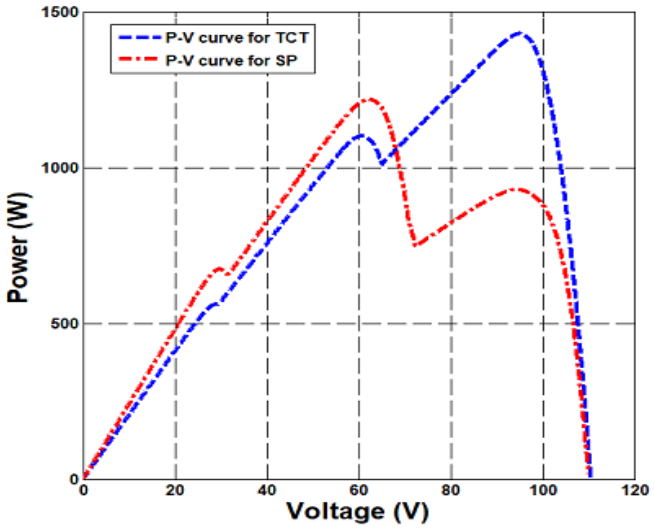


Fig. 6. *P-V* characteristics for Case 4.

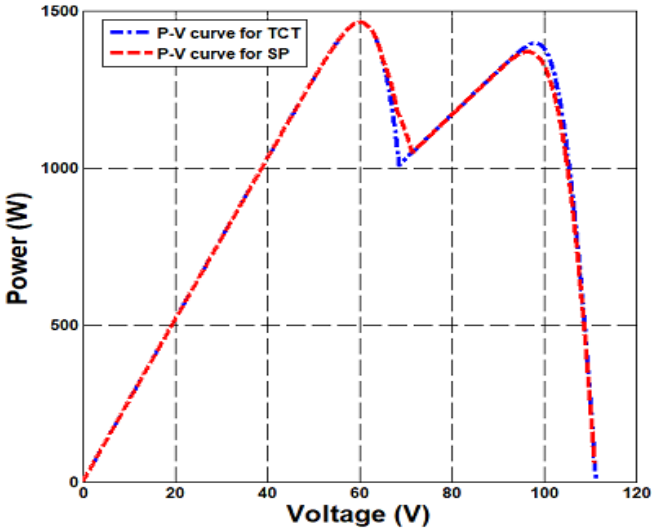


Fig. 7. *P-V* characteristics for Case 5.

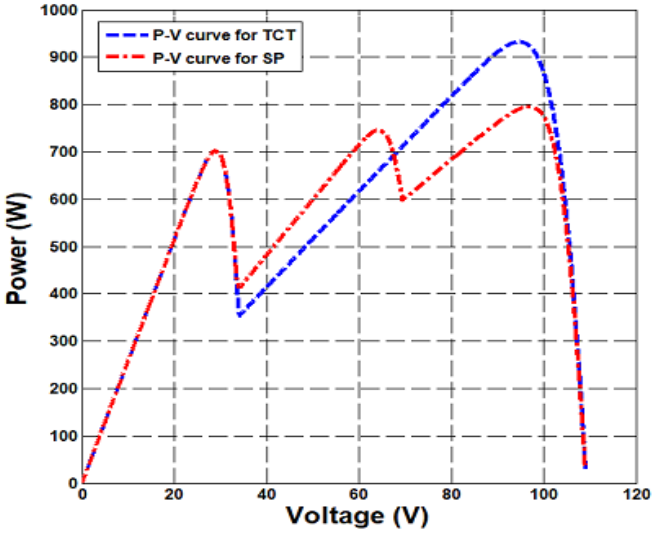


Fig. 8. *P-V* characteristics for Case 6.

The absolute values of the output parameters of SP and TCT array configuration are presented in Tables 2 and 3, respectively. It may be observed that in 2 out of 6 cases (i.e., in Cases 1 and 5), the peak power is equal for both SP and TCT arrays, while in the rest of the four cases, the power output was more in the case of TCT (also refer Fig. 9). Case 1 represents uniform irradiance over all the modules of the array and hence the similar output result; however, for Case 5, which is a horizontal shading condition, is also indifferentiable, which implies that there is no advantage of increasing the interconnection redundancy if all the modules in the same row go under shading., and TCT array will be advantageous over SP array configuration under all other shading scenarios.

Table 2. Performance parameters of SP for all the cases of shading scenarios

Cases	P_{MAX} (W)	V_{MPP} (V)	I_{MPP} (A)
Case 1	2225.4	91.78	24.24
Case 2	1637.4	92.05	17.78
Case 3	1122.6	92.78	12.09
Case 4	1219.0	61.51	19.81
Case 5	1463.3	60.34	24.25
Case 6	795.22	97.15	08.18

Table 3. Performance parameters of TCT for all the cases of shading scenarios.

Cases	P_{PEAK} (W)	V_{MPP} (V)	I_{MPP} (A)
Case 1	2225.4	91.76	24.24
Case 2	1712.2	93.22	18.36
Case 3	1290.8	92.62	13.93
Case 4	1429.2	95.63	22.67
Case 5	1463.3	60.36	24.24
Case 6	932.36	94.46	09.87

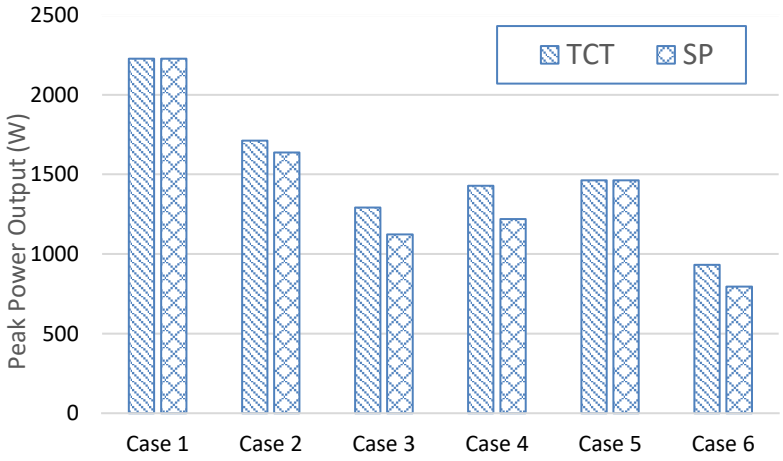


Fig. 9. Comparison of peak power outputs, under different shading scenarios.

4 Conclusions

In this paper, a comparative performance analysis of 3×3 matrix of conventional series-parallel and total cross-tied configurations is carried out under six different irradiance patterns. The produced current, voltage, and power at the global maximum power point have been reviewed to have a deeper insight into the behavior of both the PV array configurations. From the analysis, it is found that both the configurations perform equally under uniform shading and horizontal shading. Horizontal shading may be caused by adjacent rows of panels when the sun is at low altitudes, viz. during sunrise and sunsets, if the distance between each row is inadequate. TCT configuration performed better in all other shading scenarios studied in the present work. Hence there is no advantage in increasing the interconnection redundancy if the arrays are

expected to perform under uniform irradiance or self-shading under low altitudes of sun. From the present work, it is also evident that if there are possibilities of the creation of other shading patterns, such as nearby buildings, poles, trees, and non-uniform dust accumulation, the performance ratio of a PV plant may be increased by increasing the interconnection redundancy.

References

1. S. Ghosh, V. K. Yadav, G. Mehta and R. Birajdar, "Evaluation of Indian power sector reform strategies and improvement direction though DEA," 2017 IEEE Power & Energy Society General Meeting, Chicago, IL, 2017, pp. 1-5.
2. S. Ghosh, V. K. Yadav, V. Mukherjee, and S. Gupta, "Three Decades of Indian Power-Sector Reform: A Critical Assessment," *Utilities Policy*, vol. 68, pp. 101158, Feb. 2021.
3. M. K. Verma, V. Mukherjee, V. K. Yadav, and S. Ghosh, "Indian power distribution sector reforms: A critical review," *Energy Policy*, vol. 144, pp. 111672, Sept. 2020.
4. M. K. Verma, V. Mukherjee, V. K. Yadav, and S. Ghosh, "Constraints for effective distribution network expansion planning: an ample review. *International Journal of System Assurance Engineering and Management*, vol. 11, pp. 531–546, Mar. 2020. doi: 10.1007/s13198-020-00969-w.
5. S. V. Dhople, A. D. D. Garcia, A.D.D., Estimation of photovoltaic system reliability and performance metrics. IEEE Transactions on power systems, Vol. 27(1), 554-563, 2012.
6. S. Ghosh, V. K. Yadav, V. Mukherjee and R. Birajdar, "Status Check: Journey of India's Energy Sustainability through Renewable Sources," *Control of Power and Energy System, IFAC-PapersOnLine*, vol. 48 (30), pp. 456-461, 2015.
7. Palawat, K.P., Yadav, V.K., R.L. Meena and Ghosh, S. "Comparative Performance of Different PV Array Topologies under Partial Shading Condition." Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering. Lecture Notes in Electrical Engineering (Forthcoming), Springer, Singapore.
8. Palawat, K.P., Yadav, V.K., R.L. Meena and Ghosh, S. "Experimental Investigation of Performance of PV Array Topologies under Simulated PSCs," Applied Soft Computing and Embedded System Applications in Solar Energy (Forthcoming), CRC Press.
9. Yadav A.S., Yadav V.K., Mukherjee V., Ghosh S. (2021) Performance Investigation of Different Bypass Diode Topology Based SDK-PV Arrays Under Partial Shading Conditions. In: Favorskaya M., Mekhilef S., Pandey R., Singh N. (eds) Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering, vol. 661, pp. 261-270. Springer, Singapore. https://doi.org/10.1007/978-981-15-4692-1_20.
10. Gupta S., Yadav V.K., Ghosh S., Saravanan D. (2021) Non-parametric Frontier Analysis Models for Relative Performance Evaluation. In: Favorskaya M., Mekhilef S., Pandey R., Singh N. (eds) Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering, vol. 661, pp 445–454. Springer, Singapore. https://doi.org/10.1007/978-981-15-4692-1_34.
11. A. Singla, K. Singh and V. K. Yadav, "Environmental effects on performance of solar photovoltaic module," 2016 Biennial International Conference on Power and Energy Systems: Towards Sustainable Energy (PESTSE), Bengaluru, India, 2016, pp. 1-6, doi: 10.1109/PESTSE.2016.7516480.
12. S. Ghosh, V. K. Yadav and V. Mukherjee, "Impact of environmental factors on photovoltaic performance and their mitigation strategies–A holistic review," *Renewable Energy Focus*, vol. 28 (0), pp. 153-172, 2019.

13. S. Ghosh, V. K. Yadav and V. Mukherjee, "Improvement of partial shading resilience of PV array though modified bypass arrangement," *Renewable Energy*, vol. 143, pp. 1079-1093, 2019.
14. Nasiruddin, S. Khatoon, M. F. Jalil, R. C. Bansal, Shade diffusion of partial shaded PV array by using odd-even structure. Sol. Energy 181, 519-529, 2019.
15. M. G. Villalva, J. F. Gazoli, E. R. Filho, Comprehensive approach to modeling and simulation of photovoltaic arrays. IEEE Transactions on electronics, Vol. 24(5), 1198-1199, 2009.
16. D.P. Winston, G. Karthikeyan, M. Pravin, O. Jebasingh, A. G. Akash, S. Nithish, Parallel power extraction technique for maximizing the output of solar PV array. Sol. Energy 213, 102-117, 2021.
17. S. Ghosh, V. K. Yadav, V. Mukherjee, A novel hot spot mitigation circuit for improved reliability of PV module. IEEE Transactions on device and materials reliability, Vol. 20(1), 191-198, 2020.
18. Y. Yang, A. Sangwongwanich, F. Blaabjerg, Design for reliability of power electronics for grid connected photovoltaic systems. CPSS Transactions on power electronics and applications, Vol. 1(1), 92-103, 2016.
19. D. Picault, B. Raison, S. Bacha, J. Aguilera, J. D. L. Casa, changing photovoltaic array interconnections to reduce mismatch losses: a case study. EEEIC, 37-40, 2010.
20. A. Ristow, M. Begovic, A. Pregelj, A. Rohatgi, Development of a methodology for improving photovoltaic inverter reliability. IEEE Transactions on industrial electronics, Vol. 55(7), 2581 – 2592, 2008.
21. D. Wang, T. Qi Y., Liu, Y. Wang, J. Fan, Y. Wang, H. Du, A method for evaluating both shading and power generation effects of rooftop solar PV panels for different climate zones of China. Sol. Energy 205, 432-445, 2020.
22. G. Sagar, D. Pathak, P. Gaur, V. Jain, A sudoku puzzle-based shade dispersion from maximum power enhancement of partially shaded hybrid bridge-link-total-cross-tied PV array. Sol. Energy 204, 161-180, 2020.

Materials Today: Proceedings xxx (xxxx) xxx

Contents lists available at ScienceDirect

Materials Today: Proceedings

journal homepage: www.elsevier.com/locate/matpr

Plasmonic behaviour of a 1-D plasmonic nanostructure on a spacer-separated metallic film

Sreelekshmi R. Nair, Yashna Sharma *

Department of Electronics and Communication Engineering, Delhi Technological University, Shahbad Daulatpur, Delhi 110042, India

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Plasmonics
Surface plasmons
Nanowires
Finite Difference Time Domain

ABSTRACT

A plasmonic nanowire on a spacer-separated plasmonic substrate is proposed and numerically analyzed. The plasmon hybridization in the spacer region and the electric field enhancement obtained in the spacer region is modeled using Finite Difference Time Domain (FDTD) modeling. The effect of the diameter of the nanowire, the gap in the spacer region and the angle of incidence of the incoming light on the nature and intensity of electromagnetic enhancement is numerically calculated and discussed. These nanowire based systems are of prime importance for applications such as single molecule detection due to the formation of a concentrated electromagnetic hotspot in the spacer region.
© 2021 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the International Conference on Nanoelectronics, Nanophotonics, Nanomaterials, Nanobioscience & Nanotechnology.

1. Introduction

The confinement of light in nano-scale gaps by employing nanostructures has attracted large attention in the recent years as this can lead to the development of nanophotonic devices such as plasmonic nanolasers amongst others [1–4]. These ultrathin gaps lead to a significantly high electromagnetic enhancement in the gap region and support trapped plasmonic modes inside sub-wavelength volumes [5–6]. However, the fabrication of sub-wavelength gaps, especially less than 5 nm, is extremely challenging due to the limited resolution of most of the currently available nanolithography processes [7–9]. In the recent decade, the nanoparticle on metal film (NPoM) with a thin dielectric spacer has been explored for its capability to produce high electromagnetic enhancements in the spacer region. The gap plasmons created in the thin spacer region between the two metallic surfaces lead to the creation of an intense electromagnetic hotspot in these structures [10].

When nanoparticle over mirror configurations are employed, the volume of the 3-D electromagnetic hotspots over which high EM fields are present, are restricted to only nanometer-scale volumes. In this paper, we propose a plasmonic nanowire (diameter, D) over a metallic film (Fig. 1), with a dielectric spacer (thickness, t), for plasmonic enhancement. Because of the nature of the nanowire, these structures are capable of supporting a long 2-D electromagnetic hotspot along the length of the nanowire, which can offer better performance if these substrates are employed for applications such as single molecule sensing. The plasmonic properties of these structures are a strong function of the geometric parameters such as the diameter of the nanowire, the spacer thickness, the material of the nanowire and the underlying substrate, the dielectric environment, and the angle of incidence. In this paper, by employing Finite Difference Time Domain modeling, we exhaustively analyze the effect of variation in the diameter of the nanowire, the spacer thickness, and the angle of incidence on the nature of electromagnetic enhancement and the plasmon resonance wavelength achievable in these structures.

2. Numerical methods

We used Finite Difference Time Domain (FDTD) modeling to carry out the numerical simulations in this paper. FDTD uses a central approximation method to convert the continuous time Maxwell's equations into difference time equations, and solves these equations in both space and time. A commercial software (FDTD Solutions) by Lumerical Solutions, Inc was employed to analyze the electric fields in this paper. We employed 2 D FDTD modeling to simulate the various structures proposed in this paper, as the length of the nanowires was assumed to be very large compared to the wavelength of incident light. A plane wave source with an incident electric field of 1 V/m was used to illuminate the

* Corresponding author.
E-mail address: yashnasharma@dtu.ac.in (Y. Sharma).

https://doi.org/10.1016/j.matpr.2021.06.045
2214-7853/© 2021 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the International Conference on Nanoelectronics, Nanophotonics, Nanomaterials, Nanobioscience & Nanotechnology.

Please cite this article as: S.R. Nair and Y. Sharma, Plasmonic behaviour of a 1-D plasmonic nanostructure on a spacer-separated metallic film, Materials Today: Proceedings, https://doi.org/10.1016/j.matpr.2021.06.045

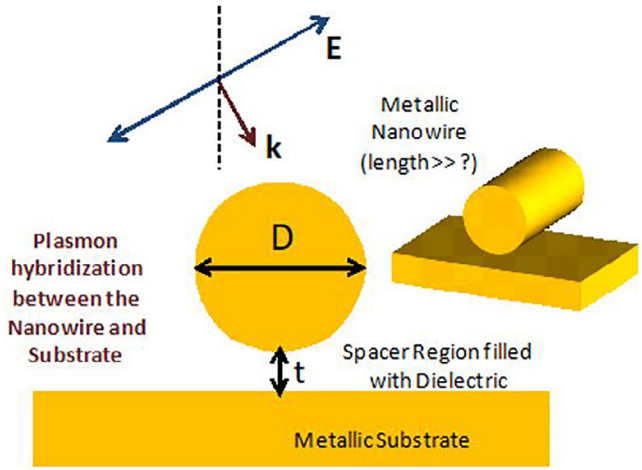


Fig. 1. Schematic of the 1-dimensional plasmonic gold nanowire (NW) on a plasmonic substrate separated by a dielectric spacer of thickness, t. Light is incident at an angle of 45 degrees with the normal, and the direction of the E-field vector is also shown.

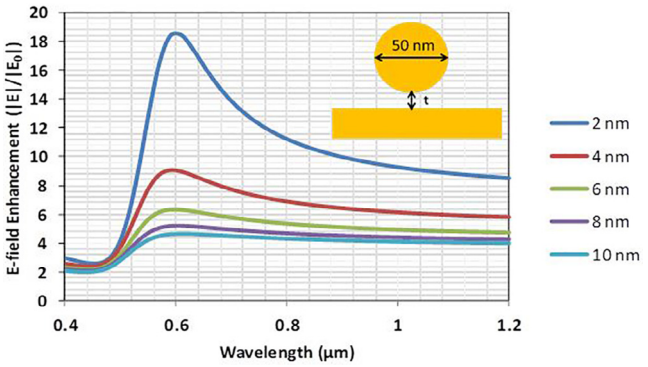


Fig. 2. Effect of the variation of the spacer thickness (t) on the E-field enhancement spectra, when the diameter of the nanowire is kept constant at 50 nm. Light is incident at an angle of 45 degrees with respect to the normal.

structure. The propagation vector (k) and the direction of polarization of the electric field vector (E) are shown in Fig. 1. The diameter of the nanowires, the thickness of the spacer layer (taken to be air in this manuscript) and the angle of incidence of illuminating light were varied to analyze their effect on the electromagnetic intensity enhancement and the plasmon resonance wavelength. Perfectly matched layer (PML) boundaries were used in the all the directions. A mesh size of 0.5 nm was employed in the x- , y- and z-directions. The dielectric constants of gold were based on the Lorentz-Drude model.

3. Results and discussions

The gold nanowire on a gold film with a thin spacer (assumed to be air in this manuscript) is numerically investigated using 2-dimensional Finite Difference Time Domain modeling. When incident light falls on the plasmonic nanowire, it excites localized surface plasmons which get coupled to the surface plasmons propagating on the gold film, leading to plasmon hybridization. Additionally, it may be noted that the direction of illumination is at 45 degrees with respect to the surface normal as shown in Fig. 1. The angle of incidence in this paper is taken to be 45 degrees, to allow both horizontal and vertical dipole modes to be excited in the nanowire and the underlying substrate.

Fig. 2 shows the effect of spacer thickness, i.e., the gap between the nanowire and the underlying substrate on the E-field enhancement spectra. For this analysis, the diameter of the nanowire was kept as 50 nm. It must be noted that the dipole charge distribution in the metallic nanowire leads to the formation of an image dipole in the substrate, and therefore the strength of the electromagnetic field enhancement is directly related to the strength of the dipole–dipole interactions in the gap. As the gap decreases, the dipole–dipole interaction increases, and leads to high electromagnetic enhancement. It can be seen in Fig. 2 that as the thickness of the spacer is decreased, the E-field enhancement increases and also experiences a slight red shift due to the increased strength of the dipole interactions.

Fig. 3 shows the E-field enhancement maps for the various values of spacer thickness starting from 2 nm to 10 nm. It can be

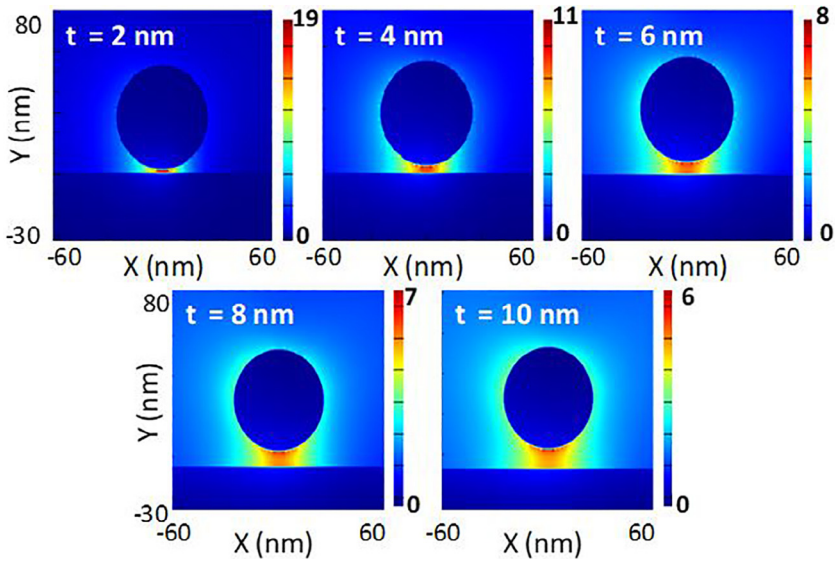


Fig. 3. E-field maps showing the E-field enhancement for different values of spacer thickness for the gold nanowire on a gold film with the thickness of the spacer being t, and the diameter of the nanowire being 50 nm. The color bars with the figures show the values of E-field enhancement, i.e., $|E|/|E_0|$ where E_0 is the incident field. Light is incident at an angle of 45 degrees with the normal.

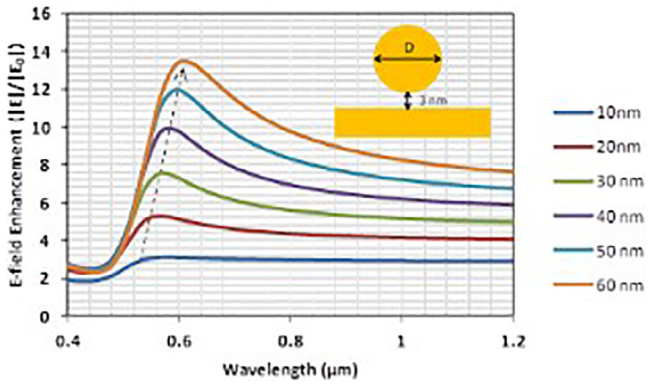


Fig. 4. Effect of the diameter of the NWs on the Electric-field enhancement spectra, when the thickness of the spacer layer is kept constant at 3 nm. Light is incident at an angle of 45 degrees with respect to the normal.

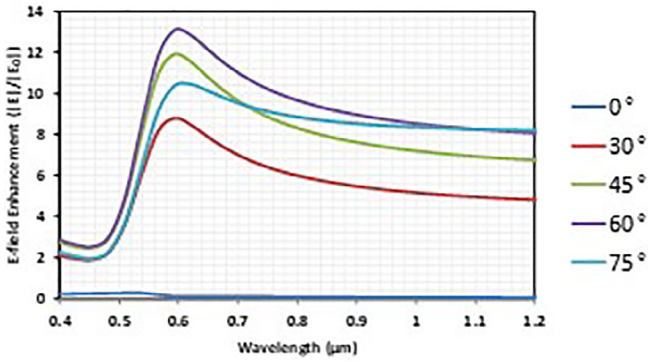


Fig. 5. Effect of variation of the angle of the incidence on the E-field enhancement spectra when the diameter of the nanowires and the thickness of the spacer layer are kept constant as 50 nm and 3 nm respectively.

observed that the confinement of the E-field is highly localized and strong when the spacer thickness is small.

Fig. 4 shows the effect of diameter on the E-field enhancement and the plasmon resonance wavelength of this structure. It can be observed that as the diameter of the nanowire increases, the E-field enhancement increases. Moreover, there is a red-shift in the plasmon resonance wavelength as the diameter of the nanowires increases. This is because the nature of localized surface plasmons on the nanowire, and the hybridized plasmons in the gap region is dependent on the diameter of the nanowire.

Fig. 5 shows the effect of varying the angle of the incidence on the nature of E-field enhancements in the proposed structure. It is clearly visible that the E-field enhancement increases as the angle increases from 0 degrees to 60 degrees as the normal component of the incident E-field increases with an increase in the angle of incidence. However, after 60 degrees, the E-field enhancement falls to lower value over a significant visible region. This is because the superposition of the incident and reflected fields is maximum at an angle of ~ 60 degrees for gold. Hence, this structure offers the maximum field enhancement when the angle of incidence is 60 degrees to the normal.

4. Conclusions

We conclude that a 1-D plasmonic nanostructure, such as the plasmonic nanowire, when placed on a spacer separated plasmonic substrate, supports hybridized plasmonic modes in the spacer region. The E-field enhancements and the plasmon resonance wavelength from these structures is a strong function of the nanowire dimensions, the nature of the substrate, the spacer thickness and the angle of incidence. Therefore, for achieving a well-localized and high E-field enhancement, the values of the spacer thickness and the diameter of the nanowire, should be appropriately chosen.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to acknowledge the Delhi Technological University for their support. The authors would also like to acknowledge the Nanophotonics and Plasmonics Lab (NPPL) at IIT Delhi for their support and resources.

References

[1] B. Lee, I.-M. Lee, S. Kim, D.-H. Oh, L. Hesselink, Review on subwavelength confinement of light with plasmonics, Journal of Modern Optics 57 (16) (2010) 1479–1497.
[2] T. Hanke, J. Cesar, V. Knittel, A. Trügler, U. Hohenester, A. Leitenstorfer, R. Bratschitsch, Tailoring spatiotemporal light confinement in single plasmonic nanoantennas, Nano letters 12 (2) (2012) 992–996.
[3] G. Toscano, S. Raza, W. Yan, C. Jeppesen, S. Xiao, M. Wubs, A.-P. Jauho, S.I. Bozhevolnyi, N.A. Mortensen, Nonlocal response in plasmonic waveguiding with extreme light confinement, Nanophotonics 2 (3) (2013) 161–166.
[4] Oulton, R. F., Sorger, V. J., Genov, D. A., Pile, D. F. P., & Zhang, X. (2008). A hybrid plasmonic waveguide for subwavelength confinement and long-range propagation. nature photonics, 2(8), 496-500.
[5] R. Chikkaraddy, X. Zheng, F. Benz, L.J. Brooks, B. de Nijs, C. Carnegie, M.-E. Kleemann, J. Mertens, R.W. Bowman, G.A.E. Vandenbosch, V.V. Moshchalkov, J. J. Baumberg, How ultranarrow gap symmetries control plasmonic nanocavity modes: from cubes to spheres in the nanoparticle-on-mirror, Acs Photonics 4 (3) (2017) 469–475.
[6] J.J. Baumberg, J. Aizpurua, M.H. Mikkelsen, D.R. Smith, Extreme nanophotonics from ultrathin metallic gaps, Nature materials 18 (7) (2019) 668–678.
[7] J.W. Thackeray, Materials challenges for sub-20-nm lithography, Journal of Micro/Nanolithography, MEMS, and MOEMS 10 (3) (2011) 033009, https://doi.org/10.1117/1.3616067.
[8] Y. Yang, C. Gu, J. Li, Sub-5 nm Metal Nanogaps: Physical Properties, Fabrication Methods, and Device Applications, Small 15 (5) (2019) 1804177, https://doi.org/10.1002/smll.v15.510.1002/smll.201804177.
[9] V.R. Manfrinato, L. Zhang, D. Su, H. Duan, R.G. Hobbs, E.A. Stach, K.K. Berggren, Resolution limits of electron-beam lithography toward the atomic scale, Nano letters 13 (4) (2013) 1555–1558.
[10] Y. Huang, L. Ma, M. Hou, J. Li, Z. Xie, Z. Zhang, Hybridized plasmon modes and near-field enhancement of metallic nanoparticle-dimer on a mirror, Scientific reports 6 (1) (2016) 1–9.

2021 2nd International Conference for Emerging Technology (INCET)
Belgaum, India. May 21-23, 2021

QnA System on Educational Textbooks: Digital Library Doubt Support System

Rahul Gupta
Department of Computer Engineering
Delhi Technological University
NCT of Delhi, India
rahulgupta100689@gmail.com

Gaurav Dabas
Department of Computer Engineering
Delhi Technological University
NCT of Delhi, India
dabasgaurav05@gmail.com

Himanshu Yadav
Department of Computer Engineering
Delhi Technological University
NCT of Delhi, India
hemanshuyadav1999@gmail.com

Nijdat Hasnain
Department of Computer Engineering
Delhi Technological University
NCT of Delhi, India
nijdathasnain1234@gmail.com

2021 2nd International Conference for Emerging Technology (INCET) | 978-1-7281-7029-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/INCET51464.2021.9456357

***Abstract* — The domain of content and knowledge on any particular topic of someone's interest is ever - growing. This broadening of the knowledge base of a subject combined with the easy accessibility to the Internet through various devices has resulted in easy access to the loads of content when a search related to the topic is made. But this gigantic collection of data also possesses a challenge. One needs to make some efforts to find the desired information. For this purpose, an individual might be required to go through a number of documents until he/she finds the answer that they craved for.**

This paper centers around a model to deal with this issue and helps in finding the relevant and precise answer to a query raised. To fulfill this purpose the model uses Natural Language Processing methods combined with knowledge graphs. The paper comes to a conclusion by presenting the performance of the model built.

Keywords — Natural Language Processing, Entity Extraction, Knowledge Graph, Parsing, Question - Answer

I. INTRODUCTION

In today's world, students are surrounded by massive chunks of information present in the form of full text documents on the Internet. But, generally a student might be more interested in knowing a direct and precise answer to his/her question rather than going through a full document to find the correct answer. All modern day search engines like Google , Microsoft Bing etc. have great capabilities but they typically display a list of documents or web pages related to the query raised by an individual. So in order to find the answer one needs to go through a complete document. Sometimes, it might also happen that a document displayed has a lot of information about the question searched but it does not include the specific information which the user is seeking.

For example,

For the following question :

"Which vitamins are present in carrots?"

Direct answers such as Vitamin A, C and K would be of greater interest to users instead of a large document which he/she needs to read to find the desired answer.

This paper facilitates finding direct answers to such questions provided the relevant text or information by using Natural Language Processing (NLP) techniques and knowledge graphs (the term made popular by Google in 2012). Since then the size of Google's knowledge graph (KG) has increased exponentially. The extent of growth has been such that Google's contains around 500 billion facts related to 5 billion entities as of May,2020. Various other organisations have also developed their knowledge graphs such as DBPedia by Wikipedia, WordNet etc. KG's have various other domain specific applications. They are used in the field of Medical Sciences, Financial Industry, Media industry and numerous other industries.

The rest of the paper is organised as follows. Section II is about related work. Section III describes the methodology used. Section IV presents our results. In section V we discuss the future scopes and limitations. Section VI lists the references.

II. RELATED WORK

A decent amount of work has been done in the past related to this field. Some of the works undertaken previously in this field are described below.

In [1] , a community QA system is deployed which when provided with multiple statements (comments) related to a question ranks them in order on the basis of good, bad and potential answers. To achieve this , the model uses convolution neural networks (CNN) with convolution filters 3,4,5. Further, long term short memory (LSTM) is used with 100 smart neurons as LSTM. Finally, the most suitable comment is provided as the answer for the question asked. This model shows an accuracy of 68.8 % with only CNN, 66.7 % with only LSTM and 72.5% with both CNN and LSTM.

Author in paper [2] has proposed a closed domain question answering system which works on education acts. Data related to education acts is collected and pre – processed using stemming and stop word removal. The keywords thus extracted are stored in an index term dictionary that contains the keyword as its key and a list of document numbers that contain the keyword corresponding to it. For a query, the keywords of the query are matched with those in the dictionary and the intersection of document numbers of all the matched keywords are retrieved and are then ranked using the jaccard similarity function. Then part of speech (POS) tagging is applied to the documents and finally the document in which the keywords have the same grammatical sense as the keywords of the query is returned as the final answer. As mentioned, this method is restricted and only returns the whole

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:52:53 UTC from IEEE Xplore. Restrictions apply.

document as an answer instead of the direct answer related to the query.

In [3] the author proposes a QA system based on DeepLSTM and versatile global T-max pooling. The model thus proposed, transforms the words in a text document to a word vector using word embedding algorithms such as word2vec. In the subsequent step the answer is predicted using a combination of deepLSTM and versatile global T-max pooling. For the purpose of ranking it uses a deep factorization classifier. Four different sources have been used for data collection such as Wikipedia. The model thus built shows an accuracy of 85.3 % on the Wikipedia dataset.

In paper [4], a method of classifying the questions according to their type has been discussed. The data set of questions used for this purpose is related to medical sciences collected from doctors and patients. It involves pre – processing the questions by stemming and removing stop words from them. Then it finds the keywords in the question and then determines the expected type of answer for that question. For example, in a question which contains the primary keyword, the answer that is expected is a "DESCRIPTION" and not a "LOCATION" or a "NUMBER". But, this model only categorises the questions and does not describe any method to predict answers for the questions.

Author in [5] proposed an intelligent QA system based on data related to basketball collected from various sources such as the NBA. Then Neo4j tool in python is used to construct a knowledge graph(KG). When a question is entered, then its similarity with the existing question in question bank is measured and if found similar, then the corresponding answer is directly returned. Otherwise, the question's entity is identified and is matched with corresponding entity in KG and all related attributes of the entity are obtained. Further, a logistic regression model is used to map the question with the attributes obtained and finally the target answer is retrieved. This system is limited only to a single entity and is not capable of finding associations in case of multiple entities.

III. METHODOLOGY

In this section we will present the overview for the data collection and its transformation which will form the basis of our structured dataset for question answering system and knowledge graph. Further we will present the algorithm for query retrieval with high precision.

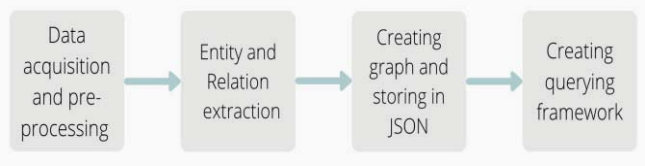


Fig. 1. Process Flow

A. Data acquisition

Even though our proposed system works on most kinds of texts but we chose our data to be textbooks of subjects like geography, history etc. in order to explain a good practical use of it.

The main source of our data is kaggle, from which we acquired some educational textbooks of the subjects of our interests (mainly theoretical subjects). Around 500 pages of each subject were considered. The data is in unstructured text format initially and not in the best form for our direct use. Thus the next step is to preprocess the data and make it suitable for our use.

B. Data Transformation

The first task is to preprocess the data to maintain uniformity in the text information. Preprocessing includes resolution of problems such as that of uppercase and lowercase in the data, conversion of pronouns into their corresponding proper nouns (also known as coherence property) to make the data more knowledgeable for the system.

The unstructured data further needs to be converted to a structured format i.e. entities and relations in order to build a knowledge graph from it and further query it too; such extracted structured data is then integrated into our knowledge graph.

C. Entity and relation extraction

We perform named entity recognition on the given data using spacy library in python. We accomplish the extraction of entities using some predefined grammar rules.

The basic idea to identify the entities is to find the subject and the object in the sentence. Similarly for relation, the verb of the sentence needs to be identified. Using parts of speech tagging we can directly find the nominal subject, direct object and the root verb in the sentence but the problem with this method is that mere nominal subject, direct object and root verb won't be enough for us to create an intelligent system. For creation of an efficient system we need to consider the whole subject, object and verb.

For example if we have a sentence, 'Indian cricketer Virat kohli is married to actress Anushka Sharma'. The nominal subject is 'Kohli', the direct object is 'Sharma' and the root verb is 'is', which is not capable of depicting any semantic knowledge. Thus, we need to consider the whole term 'Indian cricketer Virat kohli' as the subject, 'actress Anushka Sharma' as the subject and 'is married to' as the verb in order to be able to derive semantic information from these entities later. For this, we use dependency parsing along with a set of english grammar rules.

For example, to find the object of the sentence, we run a loop on the words of the sentence and check for the nsubj (nominal subject) or the nsubjpass (passive nominal subject) . Further we take in account all the compounds, nmod (nominal modifier), amod (adjectival modifier), poss (possessive nominal modifier), case (case marking) and nummods (numeric modifier) . Combining all of them in their original order will give us our required complete object. Similarly we obtain the object and the relation and we have our triplet.

We also need to take in account if any sentence has any time or place in it as it will help us to answer the questions having 'where' and 'when'. Thus to find this, for each sentence, we run a scan on the labels of each word of the sentence. If it is a date or gpe, we store it in its corresponding triplet making it a quintet. If the sentence does not have any place or time, we leave the fourth and fifth position of the quintet blank.

Quintet - [subject, predicate, object, time, place]

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:52:53 UTC from IEEE Xplore. Restrictions apply.

Now that we have a quintet for each sentence in the form of a dataframe we need to store them.

D. *Storing entities and relations in JSON format*

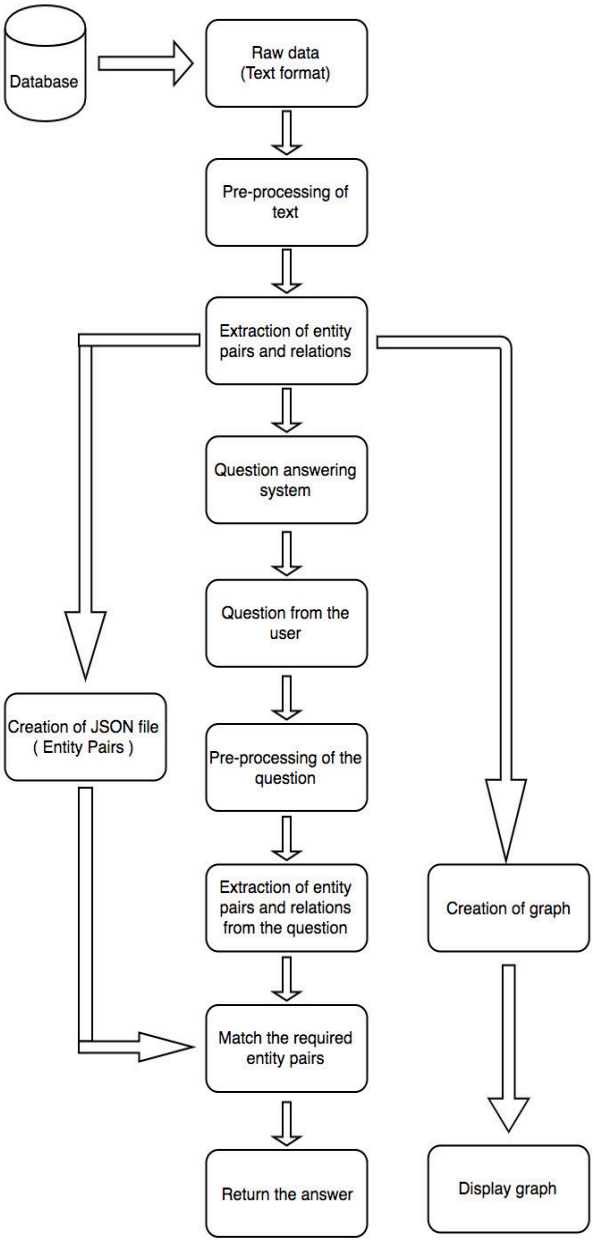


Fig. 2. Workflow chart

The derived quintets need to be stored in some structured format. Out of many format options like csv, xml, json etc. we have chosen json for the storage purpose. These quintets will be used at the time of creation of the knowledge graph and while querying the dataset.

E. *Creation of knowledge graph*

Knowledge graph is a method to display information by integrating information into an ontology and it helps the system to derive new knowledge from the graph.

It is used as it is a very flexible data structure and we can easily map our json file with the graph dynamically.

NetworkX library in python is used to display the graph. The subject, object and predicate is used from each quintet to form an edge of the graph.

F. *Query retrieval system*

The last step is to input a question from the user and answer it using our dataset. We mainly catered four types of simple questions in this system i.e. questions starting with what, who, when, where.

After preprocessing the user's question we make use of our self-defined algorithm shown below :

For questions containing **'WHO'**

```
for all entity pairs :
    if Object(Ques) == Object(entity_pair) and
        Relation(Ques) == Relation(entity_pair) :
            return Subject(entity_pair)
```

For questions containing **'WHAT'**

```
for all entity pairs :
    if Subject(Ques) == Subject(entity_pair) and
        Relation(Ques) == Relation(entity_pair) :
            return Object(entity_pair)
```

For question containing **'WHERE'**

```
for all entity pairs :
    if Subject(Ques) == Subject(entity_pair) and
        Relation(Ques) == Relation(entity_pair) and
        Object(Ques) == Object(entity_pair) :
            return Place(entity_pair)
```

For questions containing **'WHEN'**

```
for all entity pairs :
    if Subject(Ques) == Subject(entity_pair) and
        Relation(Ques) == Relation(entity_pair) and
        Object(Ques) == Object(entity_pair) :
            return Time(entity_pair)
```

Fig. 3. Pseudocode for answer extraction.

Using these rules and our previous entity and relation extraction algorithm, we derive the entity and relation from the question and return the answer if present.

IV. Result and Conclusion

For evaluation purposes we have subjected our model to a total of 500 questions out of which 450 questions have answers in the text provided and 50 do not have answers in the data. The criteria used for performance evaluation are accuracy, precision, recall and F1 score.

For calculations following attributes are used :

TP(True Positives) : total number of correct answers. Our model predicted a total of 416 correct answers out of total 450 questions whose answers are present.

FP(False Positives) : total number of incorrect answers predicted. The number of such answers in our model was 34.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:52:53 UTC from IEEE Xplore. Restrictions apply.

TN(True Negatives) : number of questions whose answers are not present in the data and hence no answer was predicted by the model. Such questions accounted to a total of 29.

FN(False Negatives) : number of questions whose answers are not present in text but the model predicted some answer. A total of 21 questions lie in this category.

Calculations are done as follows :

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$F1\ Score = \frac{2 * Precision * Recall}{Precison + Recall}$$

Table 1 displays the values of the above mentioned measures.

TABLE I. PERFORMANCE MEASURES FOR OUR MODEL.

MEASURE	Value
Precision	92.4 %
Recall	95.19 %
Accuracy	89 %
F1 Score	93.77 %

As shown in table 1 our model shows an accuracy of 89% , precision of 92.4% and recall of 95.19 % . The corresponding F1 score is 93.77 %

V. LIMITATIONS AND FUTURE SCOPE

Some of the limitations of the proposed system are :

- The rules are specified for limited sentences only, i.e. only 4 kinds of questions and have a non-complex structure.
- Redundant entities will be generated by sentences with similar meanings but distinct structures.
- extraction of entities takes into account the dependency tree of the sentence as the concept is highly dependent on parsing of the dependency tree.

Some of the future scopes of the proposed system are:

- Model can be modified to include voice recognition for the purpose of taking questions and returning answers which can be beneficial for visually impaired students.
- The algorithm of solving complex questions can be implemented by creating their own individual rules.
- The data accumulated can be trained to create neural networks for improving the QnA system.
- To boost accuracy, user reviews can be taken for accurate responses.

REFERENCES

[1] M. Wakchaure and P. Kulkarni, "A Scheme of Answer Selection In Community Question Answering Using Machine Learning Techniques," 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 2019, pp. 879-883, doi: 10.1109/ICCS45141.2019.9065834.

[2] S. P. Lende and M. M. Raghuwanshi, "Question answering system on education acts using NLP techniques," *2016 World Conference on Futuristic Trends in Research & Innovation for Social Welfare (Startup Conclave)*, Coimbatore, 2016, pp. 1-6, doi: 10.1109/STARTUP.2016.7583963.

[3] D. V. Vekariya and N. R. Limbasiya, "A Novel Approach for Semantic Similarity Measurement for High Quality Answer Selection in Question Answering using Deep Learning Methods," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2020, pp. 518-522, doi: 10.1109/ICACCS48705.2020.9074471.

[4] T. Dodiya and S. Jain, "Question classification for medical domain Question Answering system," 2016 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), Pune, 2016, pp. 204-207, doi: 10.1109/WIECON-ECE.2016.8009118.

[5] Y. Li, J. Cao and Y. Wang, "Implementation of Intelligent Question Answering System Based on Basketball Knowledge Graph," 2019 IEEE 4th Advanced Info Technology, Electronic and Automation Control Conference (IAEAC), Chengdu, China, 2019, pp. 2601-2604, doi: 10.1109/IAEAC47372.2019.8997747.

[6] D. Elalfy, W. Gad and R. Ismail, "A hybrid model to predict best answers in question answering communities", *Egypt. Inform. J.*, vol. 19, no. 1, pp. 21-31, 2018.

[7] D. Song et al., "Building and Querying an Enterprise Knowledge Graph," in IEEE Transactions on Services Computing, vol. 12, no. 3, pp. 356-369, 1 May-June 2019, doi: 10.1109/TSC.2017.2711600.

[8] Q. Shuai and C. Zhang, "Question Answering system based on Knowledge Graph of Film Culture," 2020 International Conference on Culture-oriented Science & Technology (ICCST), Beijing, China, 2020, pp. 150-153, doi: 10.1109/ICCST50977.2020.00035.

[9] Rani, K. Mehla and A. Jangra, "Parsers and parsing approaches: Classification and state of the art," 2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE), Noida, 2015, pp. 34-38, doi: 10.1109/ABLAZE.2015.7154963.

[10] T. Wang, Y. Wang and C. Tan, "Construction and Application of Knowledge Graph System in Computer Science," 2018 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC), Jinan, China, 2018, pp. 169-172, doi: 10.1109/SPAC46244.2018.8965547.

[11] Xiaohan Zou 2020 J. Phys.: Conf. Ser. 1487012016

[12] Z. Chen, Y. Wang, B. Zhao, J. Cheng, X. Zhao and Z. Duan, "Knowledge Graph Completion: A Review," in IEEE Access, vol. 8, pp. 192435-192456, 2020, doi: 10.1109/ACCESS.2020.3030076.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:52:53 UTC from IEEE Xplore. Restrictions apply.

International Journal of Environmental Science and Technology
https://doi.org/10.1007/s13762-021-03425-9

REVIEW

Relevance of phosphate solubilizing microbes in sustainable crop production: a review

F. Fatima[1] · M. M. Ahmad[1] · S. R. Verma[2] · N. Pathak[3]

Received: 15 June 2020 / Revised: 4 February 2021 / Accepted: 29 May 2021
© Islamic Azad University (IAU) 2021

Abstract
Phosphorus (P) is one of the most important factors regarding plant growth and development. Its deficiency can be compensated by adding chemical fertilizers which are expensive and have an overall negative impact on the ecosystem. Their continuous use leads to the depletion of soil fertility, accumulation of toxic elements which will affect the normal soil micro flora, leading to an imbalance in equilibrium. Phosphate solubilizing microbes (PSM) may be an alternative way to overcome the P deficiency without causing any harm to environment. For this, their biodiversity, mode of action, colonizing ability and their careful relevance should be used as trustworthy components in sustainable agricultural systems. Due to variability in the performance of these microbes, their exploitation in the area of nanotechnology becomes a challenge in sustainable agriculture. In this review, we have discussed the use of PSM as biofertilizers, responsible factors and the strains that are used for the synthesis of different nanoparticles whose implementation will not only promote sustainable agriculture but also be helpful in commercial applicability.

Keywords Crops productivity · Phosphate solubilizing microbes · Mineralization · Solubilization · Nanotechnology · Sustainable agriculture

Introduction

The universal need to increase crop yields from a slowly diminishing and degrading soil resource had drawn significant attention to the agroecosystem around the world. In the coming years, increasing crop production for the growing population in all developing countries, particularly countries like India, will be a major challenge for agro-scientists. The most popular method for improved growth is the use of chemical fertilizers, which reflects a 50 percent boost in crop production, while the other half results in increased soil alkalinity, leading to loss of soil quality and environmental degradation (Lin et al. 2019). These chemically synthesized fertilizers contain heavy acids, like nitric acid, sulfuric acid and hydrochloric acid in their composition that upon accumulation tends to reduce the soil's useful microbial population, increase soil acidity and ultimately interfere with the growth of plants. However, their continuous use may reduce the population of these beneficial microbes. Their adverse effects on the environment and increasing costs prompted the researchers to use alternate methods to increase crop production that would be environmentally sustainable and cost-effective. Implementations of plant growth-promoting microorganisms (PGPMs) may thus be seen as an effective way of minimizing the usage of chemical fertilizers within the agricultural boundaries (Dobbelaere et al. 2003).

Availability of phosphate in soil

Phosphorus (P) is an essential macronutrient known for biological and physiological development known for the cellular organization, stimulation of root development, formation of seed and in earlier maturation of crops. In nature, two forms of available P are present in the soil, *viz.*, inorganic

Communicated by Samareh Mirkia.

✉ F. Fatima
fatimafaria45@gmail.com

[1] Integral Institute of Agricultural Science and Technology, Integral University, Lucknow 226026, India

[2] Department of Biotechnology, Delhi Technological University, Delhi 110042, India

[3] Department of Biochemistry, Dr. Ram Manohar Lohia Avadh University, Faizabad 224001, India

Published online: 12 June 2021

and organic phosphates. Mostly in some types of soil only 0.1 % is utilized by the plant (Beever and Burns 1980). Still, a large proportion of phosphates remains present in unavailable form (Fig. 1). This inorganic insoluble form of P is apatite, oxyapatite and hydroxyapatite. However, under favorable conditions, they can be solubilized and become available. Depending on the soil properties, phosphate anions (PO_4^{3-}) are found to be extremely reactive and can be immobilized *via* precipitation with the Mg^{2+}, Ca^{2+}, Al^{3+}, or Fe^{3+} cations. A huge amount of P when applied as fertilizer reacts with reactive cations in acidic soil and Ca^{2+} in calcareous or normal soils *via* precipitation reaction (Richardson et al. 2011). Richardson (2011) suggested that organic matter should also exist as the main reservoir (30%–80%) of immobilized P mostly in the form of orthophosphoric acid, phytin, inositol phosphates, phospholipids, nucleic acids, etc., degraded by the activity of enzymes like acid phosphatase, phytase and nucleases.

Plant growth-promoting microorganisms (PGPM)

Agricultural activities now take more acceptable and environmentally sustainable approaches, such as the use of plant growth-promoting microbes (PGPMs), as the excessive use of chemical fertilizers and pesticides has adverse effects on the soil ecosystem as well as human health (Leach and Mumford 2011). Various important microbes have been reported to be associated within a plant's rhizospheric region which has beneficial effects on plant growth known as microbe-promoting plant growth (PGPM) (Mustafa et al. 2019). These positive effects are mediated by a series of mechanisms which includes enhancement of plant tolerance against abiotic and biotic stress, improvement in mineral nutrition, nitrogen fixation, phosphate solubilization, iron sequestration, root modification, as well as suppression of soil-borne diseases, synthesis of phytohormones, modulation of plant ethylene levels, and control of phytopathogenic microorganisms (Olanrewaju et al. 2017; Ramamoorthy et al. 2001; Ali et al. 2014; Chabot et al. 1996) as represented in Fig. 2. Microbial strains like *Pseudomonas, Rhizobium, Acetobacter, Bacillus amyloliquefaciens, Vibrio proteolyticus, Xanthobacter agilis, Enterobacter aerogenes, E. asburiae, E. taylorae, Aspergillus flavus, Trichoderma harzianum, A. niger, Penicillium rugulosum,* etc., were also identified as PGPM (Table 1). Specific bacterial rhizospheric strains such as *Rhizobium, Pseudomonas and Bacillus* are among the strongest strains with phosphorous solubilizing capacity (Sharma et al. 2013; Antoun and Pre´ Vost 2005). Kaur et al. (2015) stated that the combined inoculation of PSB with rock phosphate solubilization could serve as an alternative to the application of chemical fertilizers in sustainable agricultural systems.

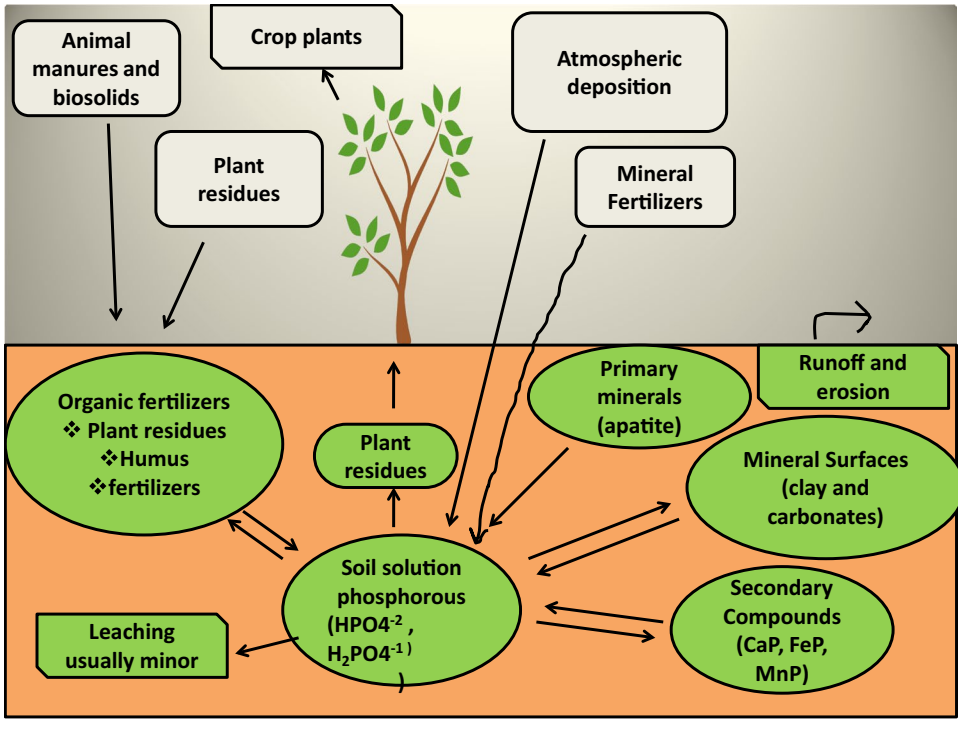


Fig. 1 Ground phosphorus balance within the soil

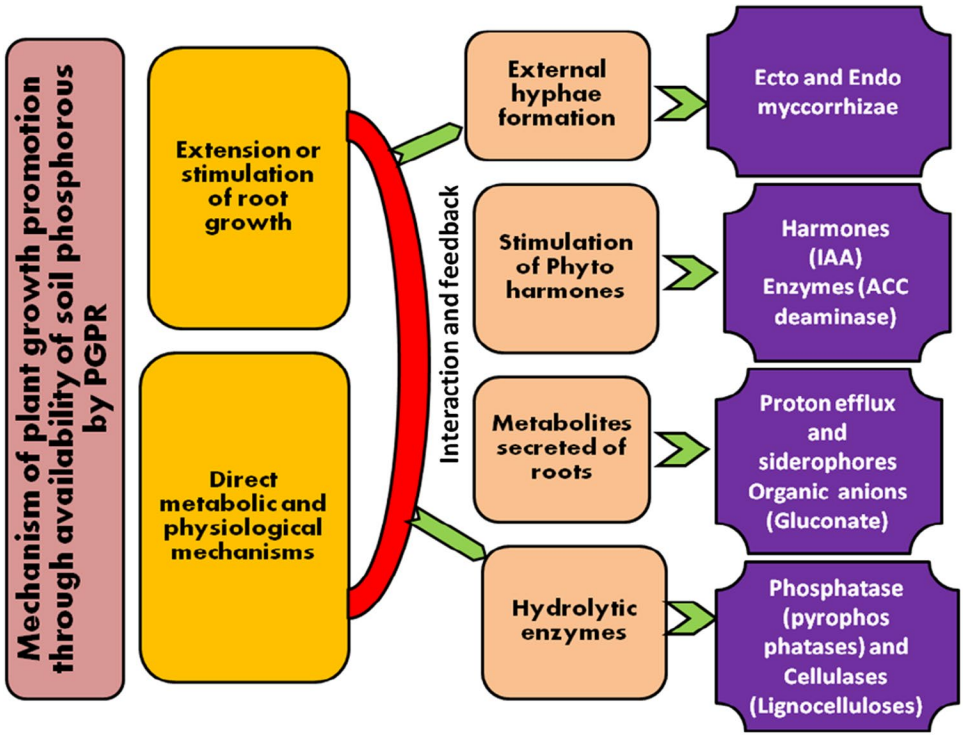


Fig. 2 Processes associated with the promotion of plant growth through microbial mediation by direct and indirect mode

Table 1 List of crops and important plant growth regulators (PGPRs)

Crops used for the study	Phosphate solubilizing bacteria as PGPR	Phytohormones secreted	References
Rice	*R. leguminosarum*	Indole-3-acetic acid	(Kim et al. 2018)
Maize	*Azotobacter sp.*	Indole-3-acetic acid	(Khalid et al. 2001)
Wheat, Brassica, Wheat	*Rhizobacterial spp.*	Auxins, IAA Acetamide	(Muratova et al. 2005); (Asghar et al. 2002)
Wheat	*Azospirillum lipoferum*	Auxin	(Ahmad et al. 2005); (Egamberdiyeva 2005)
Wheat, maize	*Pseudomonas denitrificans*	Auxin	(Ahmad et al. 2005)
Mung bean	*Pseudomonas rathonis*	Indole-3-acetic acid	(Ahmad et al. 2005)

Phosphate solubilizing microbes (PSM)

Various microbes such as bacteria, actinomycetes, fungi and algae are capable of solubilizing and mineralizing the P. A variety of rhizospheric bacteria that help to solubilize and mineralize immobilized P are *Pseudomonas spp., Bradyrhizobium, Serratia, Agrobacterium spp., Bacillus circulans*, *Rhodococcus, Sinomonas*, *Salmonella, Thiobacillus* (Chakraborty et al. 2009), *Erwinia* (Farajzadeh et al. 2012), *Azatobacter* (Kumar et al. 2012), *Bacillus spp.*, (Raza et al. 2008), *Paenibacillus* (Rashid et al. 2012), *Burkholderia* (Badawi et al. 2011), *Rhizobium* (Kuckey 1983) whereas the fungal strains carrying this ability are *Curvularia, Phoma, Alternaria, Glomus, Micromonospora, Arthrobotrys, Aspergillus, Helminthosporium, Oidiodendron, Rhizoctonia, Penicillium, Cephalosporium, Fusarium, Achrothcium, Cladosporium, Cunninghamella, Chaetomium, Mortierella, Sclerotium, Myrothecium, Paecilomyces* and many more. Kucey (1983) reported that fungus is claimed to be stronger phosphate solubilizers as compared to bacteria because they do not loss their P dissolving activity when subculturing and can traverse long distances through their hyphen thus producing more acids such as lactic acid, citric acid, gluconic acid, oxalic acid, 2-ketogluconic acid, acetic acid and tartaric acid than bacteria. Recently, actinomycetes have also attracted special attention due to their P solubilizing ability, production of phytohormones, antibiotics and are capable of surviving in extreme conditions, *e.g.*, *Actinomyces, Micromonospora* and *Streptomyces* (Sharma et al. 2013; Khan et al. 2007). It has been

documented that *Kushneria sinocarni*, a halophytic bacterium isolated from stressed environments from the Daqiao saltern east coast of China, has also shown to exhibit P solubilization properties (Zhu et al. 2011).

Mechanisms of phosphorus solubilization

A number of the mechanism has been reported for P solubilization by rhizospheric bacteria and fungi. The ability of PSM to release metabolites like low molecular weight organic acids/inorganic acids (Seshachala and Tallapragada 2012), and decrease the pH (Alori et al. 2017a). A general sketch of P solubilization is represented in Fig. 3. Besides these acid systems, phosphate deprivation also activates the phosphate solubilization system (Gyaneshwar et al. 1999). In addition, releasing root exudates such as organic ligands helps in modification of P concentration within soil (Geelhoed et al. 1999). Consequently, the process responsible for mineralizing and solubilizing organic and inorganic phosphates is discussed as: -

Solubilization and mineralization of organic P

In soil, organic P constitutes around 35–50%, which is present either in the form of inositol phosphate (soil phytate) or in phosphodiester, phospholipids, nucleic acids, pesticides, antibiotics, detergents additives and flame retardants (Singh and Satyanarayana 2011). These forms are usually resistant toward chemical hydrolysis and thus require conversion into soluble form (Pi, HPO_4^-, $H_2PO_4^-$). Halvorson et al. (1990) stated that sink phenomenon could be responsible, claiming that the value of P present in the biomass is significantly related to the decay of P present in the organic compounds. Different classes of enzymes such as phosphomonoesters (phosphatases), non-specific acid phosphatases (NSAPs) whose principal role is to dephosphorylate the phosphoanhydride or phosphodiester bond present in organic compounds.

Solubilization of inorganic phosphates

Organic acid production also plays an important role in inorganic P solubilization which involves pH reduction, enhance cation chelation, competition with P for adsorption sites or complex formation with metallic ions and insoluble P (Fe, Ca, Al). Trolove (2003) stated that the release of organic acids is represented by a direct oxidation mechanism via the pH decrease within the medium. The mechanism behind the production of these organic acids is either due to fermentation of sources presence of organic carbon (e.g., glucose, sucrose) or oxidative respiration (Kimberley et al. 2019). The organic acid formation within the rhizospheric area acidifies its surroundings that ultimately results in the release of PO_4^- ions (Goldstein and Liu 1987a). The prominent forms of acids are gluconic acid (Goldstein 1994), aspartic acid, lactic acid, tartaric acid (Bar-Yosef et al. 1999), citric acid and oxalic

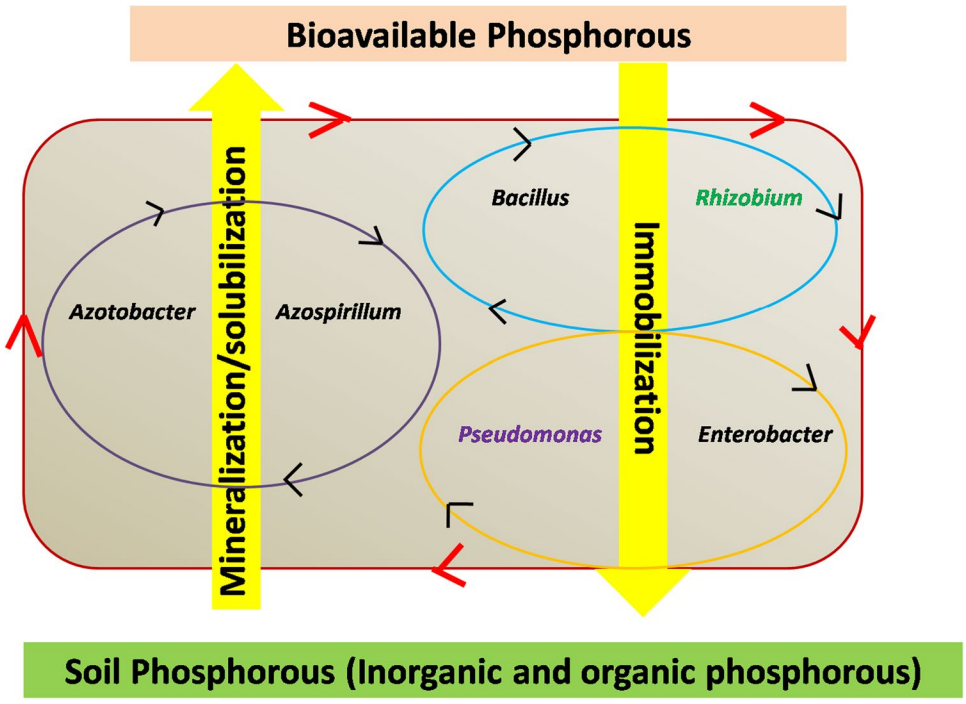


Fig. 3 Process of inorganic and organic phosphorus by solubilization and by mineralization means

acid (Venkateswarlu et al. 1984) used for solubilization. It is also reported that the ability of phosphate solubilization was done by various methods like chelation /reduction processes and H_2S gas production (Kim et al. 1997). Various studies have indicated that inorganic phosphate solubilization in the liquid media by bacterial and a fungal strain is due to carbon dioxide production. Solubilization of Fe^+ and Al^+ occurs via proton release mechanism in which adsorption of negatively charged P ions occurs. It is tested that proton release decreases P adsorption upon acidification because it increases $H_2PO_4^-$ in relation to HPO_4^{2-} which have a higher affinity toward reactive soil surfaces. Studies show that it is through rhizospheric microbes having high affinity toward iron uptake system, releasing iron-chelating molecules, *i.e.*, siderophores (Whitelaw 2000).

Genetic engineering of PSM

There is still little information about the genetics of phosphate solubilization (Altomare et al. 1999). However, certain genes involved in organic and mineral phosphate solubilization also need to be isolated and characterized. Microbial biotechnology, genetic engineering and molecular studies have created an entirely new way of modifying a number of rhizobacterial strains which improve their phosphate solubilizing ability. Insertion or over-expression of phosphate solubilizing genes in rhizospheric microbes is a novel approach for improving their ability to work as agricultural biofertilizers.

The gene cloning from the Gram-negative bacteria *Erwinia herbicola* was reported which encodes for the pyrroloquinoline quinone (PQQ) synthase gene involved in phosphate solubilization (mps). Clones containing those recombinant plasmids showed higher clearing halos with insoluble phosphate within the plates. These findings were the first study on the application of genetic modification to increase the performance of phosphate solubilization from microbial strains that can be used as biofertilizers (Goldstein and Liu 1987b). Additional studies were also carried out to investigate the function of these pqqE gene inhabiting rhizobacteria in the production of organic acid and phosphate solubilization. The pqqE successfully amplified and sequenced in genome of *Pantoea* sp. as well as *Pseudomonas* sp. Organic acids such as acetic, malic, gluconic, citric, and succinic acid were produced by these microbes. The effectiveness of phosphate solubilization was also tested by *Pantoea* sp. in the broth culture that was found as is 311 ± 4 and 204 ± 3 µg ml^{-1}, while Pseudomonas sp. showed 176 ± 3 and 298 ± 5 µg ml^{-1} of phosphate that is solubilized. The amount of acetic acid and gluconic acid was also greater in the exudates of treated plants. Maximum value was also reported for chlorophyll quality, germination rate, leaf area index, grain weight and P content in plants inoculated with these organisms. In the field trials, grain yield was improved by 10–12 % over non-inoculated control. These findings lead us to the conclusion that organic acids were produced by the rhizobacteria inhabiting the pqqE gene and phosphate was solubilized *in vitro* and *in vivo* condition (Tahir et al. 2020).

Similarly, Kim et al. (1997) reported an increase in phosphatase enzyme activity of the recombinant strain was achieved when the gene of *napA* phosphatase from the rhizospheric bacterium *Morganella morganii* was transferred into *Burkholderia cepacia IS-16* (Ishige et al. 2003). Several genes, such as *ushA* (nucleotidase), *phoA* (alkaline phosphatase), *phyC* (phytase), *pst* (Pi-specific transporter) and *glpQ* (glycerophosphoryldiester phosphodiesterase) involved in uptake of PO_4^- and solubilization of organic phosphate were stimulated during phosphate starvation (Kim et al. 1997; Ishige et al. 2003). Sequences of different phytase genes were also studied for optimal expression in *Proteobacteria*, synthesized and further engineered into the genomes of three different root-colonizing bacteria. The results showed that engineered strains showed high levels of phytate hydrolysis. Recently, *Chryseobacterium* sp., a phosphate solubilizing endophytic bacterial strain, shows plant growth promotion and biocontrol activities against *Phytophthora* blight and anthracnose on pepper using the draft genome sequence that includes genes linked to phosphate solubilization, plant growth promotion and biocontrol features (Fraga et al. 2001). These studies demonstrate that molecular approaches could be used to generate plant-associated strains with novel phosphate solubilizing capabilities (Jeong et al. 2018). Despite that, some hurdles like metabolic machinery dissimilarity and different regulating mechanism among donor and acceptor strains should be solved first to achieve successful gene introduction when using this approach. Last but not least, more studies are required to study better utilization of these PSMs under different environmental conditions and to make them favorable for farmers to boost agricultural production.

Parameters affecting phosphate solubilization

Carbon sources

The carbon source plays a very important role because of the heterotrophic nature of microorganism and of the fact that they solubilize insoluble phosphate by releasing organic acids. In the liquid medium, glucose was found to be the best energy source for microbes. PSM which can solubilize rock phosphate used various types of carbon-rich residues as sources of energy. The amount of soluble phosphate, however, varies with the form of source used

for solubilization. *Pseudomonas striata* were reported to consume glucose, galactose, sucrose and arabinose as their energy while for *Aspergillus awamori,* the preferred carbon source was sucrose, followed by arabinose, glucose, mannitol, galactose, and xylose (Shulse et al. 2019). A fungus *Penicillium digitatum, use* glucose, sucrose, mannitol, arabinose, fructose, xylose and galactose and shows maximum P solubilization efficacy.

Nitrogen sources

Microorganisms require nitrogen for the synthesis of amino acids, proteins and nitrogen bases, which may occur either in the form of ammonium salts, nitrite and nitrate, or in the form of amino acids, depending on the type of enzyme produced. Strains of *Pseudomonas striata* utilized diverse forms of nitrogen sources except for nitrites (Rudresh et al. 2004). A bacteria *Staphylococcus occidentalis* can make use of nitrogen if present either in ammonium or amino-nitrogen forms as they lack in nitrate-reducing enzymes. Ammonium sulfate and ammonium nitrate were found as the best sources among all the nitrogen sources consumed by microbial strains.

Hydrogen ion concentration

Studies show that PSM 's behavior and growth are greatly influenced by pH shift within the liquid medium. Bacteria have a maximum solubilizing pH of 5.5–6 while fungal strains have a range of 5–6. The pH activity of fungal isolates was greatly reduced at pH 7 to 8. In several studies, it was reported that at pH 4.0, fungal strains show maximum solubilization of tri-calcium phosphate.

Temperature

Temperature plays a vital role in the growth and physiological activity of PSMs. Strains of *P. striata* shows appropriate growth at a range between 25 and 30° C that while for *Aspergillus awamori* and *Penicillium digitatum*, 30°C was found as the favorable optimum temperature for P solubilization. The behaviors of certain PSMs were substantially reduced at higher temperatures (35 °C and 40 °C). Wani et al. (2007) stated that 30 °C is a suitable temperature for tricalcium phosphate solubilization by *P. digitatum*, *P. striata* and *A. Awamori* whereas bacterium *Bacillus polymyxa* can solubilize at a range between 35 and 40°C. It was also observed that the culture medium was found a less acidic at higher temperature.

Aeration

Shake culture exhibited better solubilization of rock phosphate when compared to static culture. It has been noted often that *S. occidentals* and *P. digitatum* solubilized less rock phosphate under shake culture. Several studies suggest a persistent rise in phosphate solubilization up to 9–14 days. According to several reports, after 24 hours of incubation, *Aspergillus niger* and *Pseudomonas fluorescence sp*. solubilized TCP and reached at maximum on the 7th day and 14th day (Wani et al. 2007).

Humic substances

Humic substances also have an impact on phosphate solubilization. *B. megaterium* can solubilize insoluble TCP if fulvic acid and sodium humate are present. Thus, the mechanism of phosphate solubilization is directly proportional to the number of humic substances present within the soil. Humic substances behave as strong chelating agents, and this feature of humic substances is of great use increasing the effectiveness of phosphatic fertilizers, which are rendered insoluble via fixation mechanisms (Muratova et al. 2005).

Relevance of PSM as commercial bio-fertilizer

Implementation of chemical fertilizers made farmers satisfied with rising farm yields. Yet gradually there are negative impacts such as water contamination, leaching out, fauna and flora devastation which make crops more vulnerable to microbial infections leading to a decline in soil fertility causing irreparable harm to the biosphere. Then, there comes an existence of biofertilizer, a substance containing living microorganisms which help in plant growth promotion by either increasing the availability of the nutrients or by biological control. The primary biofertilizer sources are bacteria, fungi and cyanobacteria (blue-green algae) as they have a number of capacities that could be applied to enhance agricultural practices as microbes can fight diseases, degrade/convert complex soil compounds into their simpler forms that are further used by plants to develop and grow. This proves that these microbes have a beneficial role in increasing soil fertility by producing essential nutrients for the plant and soil (Souad et al. 2018).

An alternative strategy for the application of these PSMs as inoculants is there use in either in the form of mixed cultures or as co-inoculation in combination with other beneficial microorganisms. For example, the combined effect of microbial strains (*Rhizobium*+ *Bacillus megaterium*+ *Trichoderma* spp.) on nutrient uptake growth, and yield of chickpea showed an increase in nutrient uptake, the number

of branches increased in plant height, yield of pea, nodulation and total biomass of chickpea when compared with the individual inoculations or control (Muratova et al. 2005). Co-inoculation of PSB isolates with rock phosphate promote the maize and wheat growth in terms of the height of shoot, dry biomass, the yield of grain and their total P uptake when compared with control (Quan et al. 2008).

In 2003, Nav Shakti Biotech developed a biofertilizer, Phosphobacteria which contains highly biopromoting PSB (*Bacillus megaterium*) that helps to secrete organic acids and making the unavailable phosphate into the soluble form. A biofertilizer, "Phylazonit-M" which contains a combination of three bacteria, viz., *Bacillus megaterium* + *Azotobacter chrococcum*, increases N and P supply to the plants. The primary role of these biofertilizers is to suppress the growth and activities of soil-borne pathogenic microbes in plants.

Since these microbes constitute a massive and crucial reservoir within microbiota, it can act as an innovative alternative to agricultural practices. Not only that, they also reflect the richest range of high biodiversity that helps to sustain biogeochemical cycles and food chains that can be used to grow and conserve food, manage pests, bioleaching, increase soil fertility, generate biofuels, control contaminants and serve as instruments for the therapeutic industry, such as antimicrobial agents, enzyme inhibitors, pesticides, immunomodulators/immunosuppressants and antitumor agents (Suleman et al. 2018; Fatima et al. 2015).

To solve the agricultural problems and to advance the crop biotechnology, there is an essential need to study the biodiversity of these beneficial microbes as the gene responsible for phosphate solubilization gives information related to their biochemical, evolutionary, taxonomical and metabolic and their regulatory networks. Ojuederie et al. (2017) reported that these microbes could also degrade toxic chemical pollutants and help in cleaning of polluted sites proving that these microbes act as self-sustaining and inexpensive ways. Further work is required to research the biodiversity of indigenous PSMs which can degrade various chemical contaminants and make the environment clean. Thus, more efforts have been made to characterize the microbial communities related to PSMs and their affective responses to chemical pollutants, acting as potential degraders and to identify those genes that are involved in particular degradation mechanism.

PSM performance as biocontrol agents

Phytopathogens born from soil are among the biggest issues in sustainable crop growth worldwide. To mitigate the harmful effects of pathogens on crop yields, enormous quantities of poisonous chemicals, particularly pesticides, are employed in conventional agronomic practices, which are incredibly harmful to the environment too (Ojuederie and Babalola 2017). The undesirability of implementing enormous levels of chemicals within the fields results in contamination, thus increasing resistance among soil phytopathogens. This lack of pathogen-resistant agricultural crops has compelled researchers to find alternatives to the issues with pesticides. To this end, biocontrol agents strategies consisting of biological formulations are considered the most promising choice for the use of costly destructive pesticides and the environment. In general, microorganisms with phosphate solubilizing ability were found to synthesize a broad range of metabolites with relevant biocidal and fungicidal capacities. The usage has both direct and indirect effects on plant growth and development. PSM controls pathogens by producing antimicrobial compounds such as phenolics compounds, PAL and flavonoids, antibiotics, cyanide hydrogen, indoleacetic acid (IAA), siderophores and several lytic enzymes, all of which suppress the growth of pathogens in plants (Saud and AL-Ahmadi (February 21st 2019; Köhl et al. 2019). The assessments of *Streptomyces* isolates against *B. glumae* (Zaidi et al. 2014) *Trichoderma* strains against phytopathogens (Suárez-Moreno et al. 2019) and *Trichoderma harzianum* isolate against *Sclerotinia sclerotiorum* in soybean (Bononi et al. 2020) were found as an effective biocontrol agent. However, *Burkholderia cepacia, Citrobacter freundii, Enterobacter aerogenes, Klebsiella pneumonia, Proteus vulgaris and Acinetobacter lwoffii Alternaria alternate*, *Macrophomina phaseolina*, *Fusarium oxysporum* and *Sclerotium rolfsii* were also posse's biocontrol capacity (Zhang 2016) by exerting antagonistic activities against pathogens by synthesizing antimicrobial peptides, secreting lytic enzymes and inducing systemic resistance against pathogenic microbes (Ullah et al. 2018).

Effectiveness of PSM for enhancing crop production under stress

Due to increasing populations and rising food demands, it has now become vital to mitigate the adverse soil conditions in order to increase agricultural production and, ultimately, meet the food demands. Soil and its related environmental stresses not only impact the residual microbial population species and habitats but also the productivity and yield of plants. The primary stress on the soil is metal contamination, alkalinity/salinity and drought (Kang et al. 2015). Various methods are pursued to solve this problem, and one of those solutions includes leveraging this microbial capacity for the benefit of plants. Significant agronomic features of these PSMs are to produce exopolysaccharide (EPS) (Fatima et al. 2018, 2021), various phytohormones, 1-aminocyclopropane—1 carboxylate

(ACC) deaminase, volatile compounds, enhancement of nutrient uptake, induction of accumulation of osmolytes, upregulation or downregulation of the stress responsive genes or bringing about alterations in root morphology that exaggerate the improvement in soil properties, assist in plant growth, development and biocontrol activities (Meena et al. 2017). There are multiple studies on the enhancement of microbial-mediated plant growth under stress-free conditions while fewer studies are available on their impact under stressful environment.

It demonstrates the microbial ability to improve plant tolerance in stressed soils, and their effects as a mixture on various crops. The key idea here is to integrate the fact that PSM alliance can be used primarily in stressed soil, with the goal of rehabilitating soil conditions to promote crop productivity (Köhl et al. 2019). *Bacillus*, *Pseudomonas* spp. (Panwar et al. 2014) and *Providencia rettgeri* sp. *Bradyrhizobium japonicum*, (Jiang et al. 2018), *Bacillus amyloliquefaciens* (Han and Lee 2005), *Bacillus megaterium* (Shao et al. 2015) and *Rhizobium* (Thant and Aung 2018) were able to possess strong capacity for dissolving calcium phosphate under saline conditions *Bacillus halodenitrificans* and *Hallobacillus* sp. also showed more than 90 % higher root elongation and 17.4 % higher dry weight than untreated control wheat seedlings at 320 mM NaCl stress that suggests a substantial reduction in the detrimental effects of NaCl (Surange et al. 1997). The bacterial species of *Pseudomonas fluorescens* and *Azospirillum brasilense* also showed auxin production and inorganic phosphate solubilization under hypersaline conditions (Ramadoss et al. 2013), *Pseudoalteromonas* sp. showed halo-alkali and thermostable properties proved to have potential industrial activity (Kadmiri et al. 2018). However, *Talaromyces aurantiacus* and *Aspergillus neoniger* could survive in acidic environments and both fungi had a considerable ability to release soluble P by decomposing recalcitrant P-bearing compounds in alkaline conditions (Makhdoumi et al. 2015). Same results were found in the case of *A. niger* (Zhang et al. 2018) and *A. terreus* strains that had the potential to deal with salinity and drought stress (Li et al. 2015).

Interaction of exogenous phosphorus-solubilizing bacteria with rhizospheric microbial community

Phosphate solubilizing bacteria have made substantial use in the form of biological fertilizers (PSB). However, less understanding of its effect on the general microbial ecosystem has been shown. In several studies, a mixture of PSB was inoculated in conjunction with the manure fertilizer in tomato plant. The growth parameter findings showed that the mixture of PSB and compost could be significantly induced the growth of tomato. Adding PSB could significantly increase pH, available phosphate and many types of trace minerals, both in the rhizosphere and non-rhizosphere soil. The results of quantitative PCR and high-throughput sequencing showed that the rhizosphere-inoculated PSB did not become the dominant strain. However, the structure of the soil microbial species gets modified. Some indigenous bacteria, such as Pseudomonas, have declined in number, while several bacteria have expanded in the population, including members of *Anaerolineaceae, Bacillus, Gemmationadaceae and Cytophagaceae*. The PSB secreted low-molecular organic compounds and changed soil properties in order to dissolve inorganic phosphorus, which indirectly altered the root system's microbiota community. The results of the durability analysis showed that the soil's properties had a significant effect on the endogenous bacterial population (Liu et al. 2020).

Furthermore, in phosphorus poor soils or soils with rock phosphate, PSM associates well with arbuscular mycorrhizal fungi. Certain phosphatic ions that are accessed and transferred by the arbuscular mycorrhizal fungal hyphae to the plant may be produced by PSM from some of the more sparingly soluble phosphorus reservoirs. In addition, compared to non-mycorrhizal roots, PSM persists better around mycorrhizal roots and works symbiotically with AMF, contributing to increased plant growth, particularly where rock phosphate is added to the soil. In some cases where the soil is phosphate, combined dual inoculation of AMF and PSM has been shown to growth of plants more than inoculation of either microorganism individually. The simpler understanding of this finding is that in abundance, success and survival, the mycorrhizal endophyte may be stimulated. Increased nutrient absorption, especially phosphorus, is the main effect of this mycorrhiza in enhancing plant growth due to the discovery of the existing hyphae of the soil far beyond root-hair region where phosphorus gets reduced. In addition, by PSM inoculation, root exudation and plasticity can alter, which could also influence the growth of AMF. In the root region, the AMF also produces plant hormones and enhances the development of nitrogen-fixing species. However, for the optimum use of these interactions with regard to growth and development, a more comprehensive evaluation of the dynamics among soil microorganisms is required (Alori et al. 2017b).

Canavalia ensiformis, a forage legume popularly used as organic fertilizer, similar studies have been reported. An agronomically effective and sustainable solution applied to economically important crops would be the use of AMF and phosphorus solubilizers. Mycorrhizal colonization, phosphorus availability and ultimate phosphorus solubilizers will be promoted by the interaction of AMF and phosphorus solubilizers. An improvement of 3.07 mg/Kg of soluble P in the soil favored this interaction. The number of nodules varied

from 46 to 76. The use of interactions between AMF and PSF increased mycorrhizal colonization, soil phosphorus availability and C growth (Padrón-Rodríguez et al. 2020). AMF and PSB combine synergistically when PSB solubilizes orthophosphate into sparingly accessible phosphorous compounds that can be consumed and transferred to the host plant by AMF. In conjunction with rock phosphate, the co-inoculation of KV with RI showed the biggest improvements in growing plants and tuber inulin quality relative to both unfertilized and fertilized controls. This strong cumulative effect may have had the opportunity for prospective farmers to use this association as a biofertilizer for inulin production in *Helianthus tuberosus* L (Nacoon et al. 2020).

In Chickpea (Cicer arietinum) several bio-inoculants are encouraged as natural "nutrient" resource. Bio-inoculants diazotroph *Pseudomonas palleroniana* and a phosphate solubilizer strain of *Pseudomonas jesenii* were evaluated for their plant growth-promoting potential. Furthermore, denaturing gradient gel electrophoresis (DGGE) and real-time PCR (qPCR) techniques were also implemented to analyze the microbial diversity and dynamics. Comparative agronomical and biochemical crop parameters revealed that organic soils co-inoculated with bio-inoculants and phosphate solubilizers have improvement in morphological and physiological content, respectively, over the control. Furthermore, microbial soil dynamics based on qPCR and DGGE showed the survival of strains until around the end of harvesting time without disrupting the indigenous microflora. Conclusively, their co-inoculation has demonstrated considerable promise for the growth and production of chickpeas under organically controlled soils and would also offer further motivation for their bio-formulation to progress toward zero budget natural farming (Rawat et al. 2019).

Performance of PSM in biosynthesis of metallic nanoparticles

As a strong technology, nanoscience enables us to look at the atomic and molecular levels and build structures on a nanometer scale. Nanotechnology offers agricultural productivity for the larger population of agriculture and food processing, and it promotes and frames the next stage of growth of genetically engineered plants, sustainable agriculture inputs, agricultural chemicals and advanced farming techniques of research and development (Jalili et al. 2020). In recent times, the nanomaterials have been developed using physical and chemical approaches. These methods are expensive and use harmful chemicals under ultraviolet irradiation, laser ablation and aerosol spray during the synthesis of nanoparticles. While such techniques are widely used, there is concern over the use of hazardous chemicals. (Prasad et al. 2017). Microbial synthesis of metallic nanoparticles is being investigated to overcome this problem. The physical solution of lower synthesis limits for the production of nanoparticles needs greater energy usage, and expenses are indeed very high. Industrial solutions, however, are low cost but include the use of hazardous solvents and a trace of chemical waste with hazardous by-product production (Thakkar et al. 2010). Biological approach that includes bacteria, fungi, yeast and viruses therefore provides high production, low-cost and environmentally safe nanoparticles synthesis (Narayanan and Sakthivel 2010).

Microbial nanoparticle synthesis would be a green chemistry method and is regarded an extremely interesting field for potential applications (Mandal 2006). Microbial production with enriched microbial biodiversity, using specific microbial cultivation, can be accomplished under cellular, biochemical and molecular processes, the rate of synthesis and improvement of nanoparticular properties. The relationship between microbial and metals is well recognized in biotechnological applications such as bioleaching, biomineralization, biocorrosion and bioremediation (Mohanpuria et al. 2008). The properties of microbially generated nanostructured mineral particles and metallic nanoparticles are close to those of chemically synthesized nanomaterials. Most unicellular and multicellular organisms, such as bacteria (prokaryotes) and fungi (eukaryotes), make inorganic materials either intracellular or extracellular (Mandal 2006; Gericke and Pinches 2006a). The intracellular processes include the movement of positive ions through the cell wall by contact with negative cell membrane ions. Furthermore, cell wall enzymes reduce the metal ion to nanoparticles, and later, these nanoparticles spread across the bacterial cell wall. The extracellular generation of nanoparticles in the case of fungi is the nitrate reductase-mediated production in the presence of enzyme nitrate, which helps to bio-reduce metal ions into nanoparticles (Duygu et al. 2019; Hulkoti and Taranath 2014). It provides a short description of recent PSM work involving in metallic nanoparticles synthesis.

The synthesis method (GS) of various metallic nanoparticles (MNPs) has reassessed microorganisms for their universal ability to turn heavy metal ions into neutral atoms at no cost to toxic chemicals. Microbial-mediated biosynthesis provides advantages of improved biocompatibility as compared to chemically synthesized MNPs and hence has greater potential for biological as well as agricultural applications. Microbial-mediated biosynthesis includes microorganisms such as fungi, bacteria and algae. They produce MNPs such as gold (Au), silver (Ag), zinc oxide (ZnO), iron (Fe_2O_3), nickel oxide (NiO) palladium (Pd), copper oxide (CuO), platinum (Pt) and magnesium oxide (MgO) that play a significant role as nanofertilizer (Mandal 2006), nanopesticides, nondiagnostic agents (Fatima et al. 2021), nano-induced resistance, etc. (Fatima et al. 2020). The tolerance of bacteria toward metal ions and formation

of metal nanoparticles depends on physical and chemical growth parameters, such as particle size, pH, temperature, culture time, composition of growth medium and growth in light or dark conditions. PSM and different fungal species have also shown their abilities in the reduction of metal ions. However, not a thumb rule, most of the reducing agents sourced from living world also act as capping agents and render MNPs less toxic or more biocompatible. Various examples are given in Table 2.

Thus, in the field of agriculture, nanotechnology has been used to heighten the crop production with quality enrichment by improving farming systems. The emergence of biologically synthesized nanomaterials by PSM and their actions within the frame of sustainable agriculture have revolutionized world agriculture canvass dramatically by novelty, fast growth and enormity to meet the projection of global food demand (Singh and Prasad 2017). Therefore, further research is urgently needed to unravel the behavior and fate of these PSM in altered agriculture inputs and their interaction with biomacromolecules present in living systems and environments.

Conclusion

The inoculation of PSM in soil indicates to be a successful way to turn the insoluble P compounds into the plant-available P type, resulting in improved plant growth, crop production and performance. The most important P solubilizers for increasing the bioavailability of P in the soil are *Bacillus, Pseudomonas, Rhizobium, Aspergillus, Penicillium,* and *AMF*. By supplying readily absorbable P type, manufacturing plant hormones and antibiotics such as IAA and GA, PSM causes immediate crop production. PSM promotes plant growth by siderophore development and improves nitrogen fixation ability. In addition, through the synthesis of antibiotics, hydrogen cyanate (HCN), and antifungal metabolites, PSM acts as a disease control toward pathogenic bacteria. There are also possible replacements for inorganic phosphate fertilizers in order to satisfy plant P specifications, improving yields in sustainable farming. An economical and environmentally sustainable approach is their implementation. Further analysis, therefore, is critical in the field study to explore productive biofertilizers-PSM with several growth-stimulating attributes. The diverse form of PSMs has proved an important functional activity in agricultural sector as they constitute an important role by recycling the nutrients, pollutants degraders, waste treatment, a biocontrol agent and as an important source for the production of nanoparticles in the form of nanofertilizers that boost the agricultural crop production in the coming era. However, still there is an urgent need to study these environment-friendly microbes that can solve and show a new way to advance biotechnology. Despite knowing the significance of phosphate solubilizing microbes, still the information related to their diversity

Table 2 Role of PSM in production of different nanoparticles

Microbial species	Type of NP	Size of NPs	References
Bacillus subtilis	Au	5–25 nm, octahedral	(Egamberdiyeva 2005)
Lactobacillus sp.	Au, Ag, Au–Ag, Ti	20–60 nm, hexagonal/contour/ Spherical	(Beveridge and Murray 1980)
Pseudomonas aeruginosa JP-11	CdS	5–200 nm	(Nair and Pradeep 2002)
Shewanella loihica PV-4	Pt	30–0.3 nm, spherical, chains, dendritic	(Chakraborty et al. 2018)
Pseudomonas stutzeri AG259	Ag, Ag2S	< 200 nm, nano-crystals	(Ahmed et al. 2018)
Fusarium oxysporum	Au	20–40 nm spherical, TRIANGULAR	(Naik et al. 2017)
F. oxysporum	Zr	3–11 nm quasi-spherical	(Mukherjee et al. 2002)
F. oxysporum	Au–Ag	8–14 nm	(Ishida et al. 2014)
F. oxysporum	Pt	10–50 triangle, hexagons, square, rectangles	(Senapati et al. 2005)
Verticillium sp.	Ag	25 spherical	(Syed and Ahmad 2012)
V. luteoalbum	Au	< 10 nm spherical	(Mukherjee et al. 2001)
A. flavus	Ag	8.9 nm	(Nicomrat et al. 2016)
Coriolus versicolor	Ag	25–75 nm, 444–491 nm spherical	(Gericke and Pinches 2006b)
Trichoderma viride	Ag	Spherical	(Vigneshwaran 2007)
A. fumigatus	Zn	60–80 nm spherical	(Fayaz et al. 2009); (Srivastava et al. 2016)
Aspergillus oryzae	$FeCl_3$	10–24.6	(Raliya 2013)
Aspergillus flavus TFR7	TiO_2	12–15	(Raliya et al. 2015)
Aspergillus flavus	Ag	50 nm	(Raliya et al. 2015);(Fatima et al. 2016); (Chaudhary et al. 2020)

is very insufficient. Thus, their evaluation, exploration and exploitation are needed for scientific purposes as well as further work is urgently required to unravel the actions and fate of these PSMs in altered inputs from agriculture and their association with biomacromolecules present in living systems and ecosystems for the betterment of humankind.

Acknowledgements The authors are highly thankful to Chancellor, Integral University for his support and encouragement.

Funding No funding is received to carry out this work.

Declarations

Conflict of interest The authors declares that they have no competing of interest.

References

Ahmad F, Ahmad I, Khan MS (2005) Indole acetic acid production by the indigenous isolates of *azotobacter* and flourescent *Pseudomonas* in the presence and absence of tryptophan. Turk J Biol 29:29–34

Ahmed E, Kalathil S, Shi L, Alharbi O, Wang P (2018) Synthesis of ultra-small platinum, palladium and gold nanoparticles by *Shewanella loihica* PV-4 electrochemically active biofilms and their enhanced catalytic activities. J Saudi Chem Soc 22:919–929

Ali S, Charles TC, Glick BR (2014) Amelioration of high salinity stress damage by plant growth-promoting bacterial endophytes that contain ACC deaminase. Plant Physiol Biochem 80:160–167

Alori ET, Glick BR, Babalola OO (2017) Microbial phosphorus solubilization and its potential for use in sustainable agriculture. Front Microbiol. https://doi.org/10.3389/fmicb.2017.00971

Alori ET, Glick BR, Babalola OO (2017a) Microbial phosphorus solubilization and its potential for use in sustainable agriculture. Front Microbiol 8:971

Altomare C, Norvell WA, Björkman T, Harman GE (1999) Solubilization of phosphates and micronutrients by the plant growth promoting and biocontrol fungus *trichoderma harzianum* rifai 1295–22. Appl Environ Microbiol 65:2926–2933

Antoun, H., Pre´ Vost, D., (2005) Ecology of Plant Growth Promoting Rhizobacteria, PGPR: Biocontrol and Bio fertilization. Springer, Dordrecht. 1–38

Asghar HN, Zahir ZA, Arshad M, Khaliq A (2002) Relationship between in vitro production of auxins by rhizobacteria and their growth-promoting activities in *brassica junceal*. Biol Fertil Soils 35:1–237

Badawi FSF, Amm B, Desoky AH (2011) Peanut plant growth and yield as influenced by co-inoculation with *bradyrhizobium* and some rhizomicroorganisms under sandy loam soil conditions. Ann Agric Sci 56:17–25

Bar-Yosef B, Rogers RD, Wolfram JH, Richman E (1999) *Pseudomonas cepacia*-mediated rock phosphate solubilization in kaolinite and montmorillonite suspensions. Soil Sci Soc Am J 63:1703–1708

Beever RE, Burns (1980) Phosphorus uptake, storage and utilization by fungi, advances in botanical. Research 8:127–219

Beveridge T, Murray R (1980) Sites of metal deposition in the cell wall of *Bacillus subtilis*. J Bacteriol 141:876–887

Bononi L, Chiaramonte JB, Pansa CC (2020) Phosphorus-solubilizing Trichoderma spp from Amazon soils improve soybean plant growth. Sci Rep 10:2858

Chabot R, Antoun H, Cescas MP (1996) Growth promotion of maize and lettuce by phosphate solubilizing *rhizobium leguminosarum* biovar phaseoli. Plant Soil 184:311–321

Chakraborty U, Chakraborty BN, Basnet M, Chakraborty AP (2009) Evaluation of *ochrobactrum anthropi* trs-2 and its talc based formulation for enhancement of growth of tea plants and management of brown root rot disease. J Appl Microbiol 107:625–634

Chakraborty J, Mallick S, Raj R (2018) Functionalization of extracellular polymers of *pseudomonas aeruginosa* n6p6 for synthesis of cds nanoparticles and cadmium bioadsorption. J Polym Environ 26:3097–3108

Chaudhary P, Fatima F, Kumar A (2020) Relevance of nanomaterials in food packaging and its advanced future prospects. J Inorg Organomet Polym 30:5180–5192

Dobbelaere S, Vanderleyden J, Okon Y (2003) Plant growth-promoting effects of diazotrophs in the rhizosphere. CRC Crit Rev Plant Sci 22:107–149

Duygu DY, Erkaya IA, Erdem B (2019) Characterization of silver nanoparticle produced by *Pseudopediastrum boryanum* (Turpin) E. Hegewald and its antimicrobial effects on some pathogens. Int J Environ Sci Technol 16:7093–7102

Egamberdiyeva D (2005) Plant growth promoting rhizobacteria isolated from a calsisol in semi arid region of uzbekistan: biochemical characterization and effectiveness. J Plant Nutr Soil Sci 168:94–99

Farajzadeh D, Yakhchali B, Aliasgharzad N, Bashir NS, Farajzadeh M (2012) Plant growth promoting characterization of indigenous *azotobacteria* isolated from soils in Iran. Curr Microbiol 64:397–403

Fatima F, Bajpai P, Pathak N, Singh S, Priya S, Verma SR (2015) Antimicrobial and immunomodulatory efficacy of extracellularly synthesized silver and gold nanoparticles by a novel phosphate solubilizing fungus *Bipolaris tetramera*. BMC Microbiol 15:52

Fatima F, Verma SR, Bajpai P, Pathak N (2016) Extracellular mycosynthesis of silver nanoparticles and their microbicidal activity. J Glob Antimicrob Re 7:88–92

Fatima F, Pathak N, Verma SR, Bajpai P (2018) Toxicity and immunomodulatory efficacy of biosynthesized silver myconanosomes on pathogenic microbes and macrophage cells. Artif Cells Nanomed Biotechnol 46:1637–1645

Fatima F, Siddiqui S, Khan WA (2020) Nanoparticles as novel emerging therapeutic antibacterial agents in the antibiotics resistant era. Biol Trace Elem Res 2020:1–13

Fatima F, Hashim A, Anees S (2021) Efficacy of nanoparticles as nanofertilizer production: A review. Environ Sci Pollut Res Int 28:1292–1303

Fayaz AM, Balaji K, Kalaichelvan PT (2009) Venkatesan R. Fungal based synthesis of silver nanoparticles—An effect of temperature on the size of particles. Colloids Surf B Biointerfaces 74:123–126

Fraga R, Rodriguez H, Gonzalez T (2001) Transfer of the gene encoding the Nap A acid phosphatase from Morganella morganii to a Burkholderia cepacia strain. Acta Biotechnol 21:359–369

Geelhoed JS, Riemsdijk WHVAN, Findenegg GR (1999) Simulation of the effect of citrate exudation from roots on the plant availability of phosphate adsorbed on goethite. Eur J Soil Sci 50:379–390

Gericke M, Pinches A (2006a) Biological synthesis of metal nanoparticles. Hydrometallurgy 83:132–140

Gericke M, Pinches A (2006b) Microbial production of gold nanoparticles. Gold Bull 39:22–28

Goldstein A. H: (1994). Involvement of the Quinoprotein Glucose Dehydrogenase In The Solubilization Of Exogenous Mineral Phosphates By Gram-Negative Bacteria, Phosphate In

Microorganisms: Cellular And Molecular Biology, Washington, DC, 197–203
Goldstein AH, Liu ST (1987a) Molecular cloning and regulation of a mineral phosphate solubilizing gene from *erwinia herbicola*. Biotechnology 5:72–74
Gyaneshwar P, Pareka LJ, Archana G, Podle PS, Collins MD, Huston RA, Naresh KG (1999) Involvement of a phosphate starvation inducible glucose dehydrogenase in soil phosphate solubilization by *enterobacter asburiae*. FEMS Microbiol Lett 171:223–229
Halvorson HO, Keynan A, Kornberg HL (1990) Utilization of calcium phosphates for microbial growth at alkaline phosphate. Soil Biol Biochem 22:887–890
Han HS, Lee KD (2005) Physiological responses of soybean—inoculation of *Bradyrhizobium japonicum* with PGPR in saline soil conditions. Res J Agric Biol Sci 1:216–221
Hulkoti NI, Taranath T (2014) Biosynthesis of nanoparticles using microbes—A review. Colloids Surf B 121:474–483
Ishida K, Cipriano TF, Rocha GM, Weissmüller G, Gomes F, Miranda K, Rozental S (2014) Silver nanoparticle production by the fungus *Fusarium oxysporum*: nanoparticle characterization and analysis of antifungal activity against pathogenic yeasts. Mem Inst Oswaldo Cruz 109:220–228
Ishige T, Krause M, Bott M, Wendisch VF, Sahm H (2003) The phosphate starvation stimulon of *Corynebacterium glutamicum* determined by DNA microarray analy ses. J Bacteriol 185:4519–4529
Jalili B, Bagheri H, Azadi S (2020) Identification and salt tolerance evaluation of endophyte fungi isolates from halophyte plants. Int J Environ Sci Technol 17:3459–3466
Jeong JJ, Sang MK, Pathiraja D, Park B, Choi IG, Kim KD (2018) Draft genome sequence of phosphate-solubilizing *Chryseobacterium* sp. strain ISE14 biocontrol plant growth promoting rhizobacterium isolated from cucumber. Prokaryotes 6:12–18
Jiang H, Qi P, Wang T, Wang M, Chen M, Chen N, Pan L, Chi X (2018) Isolation and characterization of halotolerant phosphate-solubilizing microorganisms from saline soils. Biotech 8:461
Kadmiri IM, Chaouqui L, Azaroual SE (2018) Phosphate-solubilizing and auxin-producing rhizobacteria promote plant growth under saline conditions. Arab J Sci Eng 43:3403–3415
Kang SM, Radhakrishnan R, Lee IJ (2015) *Bacillus amyloliquefaciens* subsp. plantarum GR53, a potent biocontrol agent resists *Rhizoctonia* disease on Chinese cabbage through hormonal and antioxidants regulation. World J Microb Biot 31:1517–1527
Kaur G, Reddy M (2015) Effects of phosphate-solubilizing bacteria, rock phosphate and chemical fertilizers on maize-wheat cropping cycle and economics. Pedosphere 25:428–437
Khalid A, Arshad M, Zahir ZA (2001) Factor affecting auxin biosynthesis by wheat a rice rhizobacteria. Pak J Soil Sci 21:11–18
Khan MS, Zaidi A, Wani PA (2007) Role of phosphate-solubilizing microorganisms in sustainable agriculture - a review. Agron Sustain Dev 27:29–43
Kim KY, Mcdonald GA, Jordan D (1997) Solubilization of hydroxyapatite by *enterobacter agglomerans* and cloned *escherichia coli* in culture medium. Bio Fertil Soils 24:347–352
Kim DY, Kadam A, Shinde S, Saratale RG, Patra J, Ghodake G (2018) Recent developments in nanotechnology transforming the agricultural sector: a transition replete with opportunities. J Sci Food Agric 98:849–864
Kimberley D, Schneider JR, Thiessen M, Francis Z, Keith D, Tandra R, Derek DF, Lynch H, O'Halloran IP, Henry F (2019) Options for improved phosphorus cycling and use in agriculture at the field and regional scales. J Environ Quality 48(5):1247–1264
Köhl J, Kolnaar R, Ravensberg WJ (2019) Mode of action of microbial biological control agents against plant diseases: relevance beyond efficacy. Front Plant Sci 10:845
Kuckey (1983) Phosphate-solubilizing bacteria and fungi in various cultivated and virgin alberta soils. Canad J Soil Sci 63:671–678
Kumar P, Dubey RC, Maheshwari DK (2012) *Bacillus* strains isolated from rhizosphere showed plant growth promoting and antagonistic activity against phytopathogens. Microbiol Res 167:493–499
Leach AW, Mumford JD (2011) Pesticide environmental accounting: a decision-making tool estimating external costs of pesticides. J Verbrauch Lebensm 6:21–26
Li X, Luo L, Yang J, Li B, Yuan H (2015) Mechanisms for solubilization of various insoluble phosphates and activation of immobilized phosphates in different soils by an efficient and salinity-tolerant *Aspergillus niger* strain An2. Appl Biochem Biotech 175:2755–2768
Lin W, Lin M, Zhou H, Wu H, Li Z, Lin W (2019) The effects of chemical and organic fertilizer usage on rhizosphere soil in tea orchards. PLoS ONE 14:e0217018
Liu J, Qi W, Li Q (2020) Exogenous phosphorus-solubilizing bacteria changed the rhizosphere microbial community indirectly. Biotech 10:164
Makhdoumi A, Dehghani-Joybari Z, Mashreghi M (2015) A novel halo-alkali-tolerant and thermo-tolerant chitinase from *Pseudoalteromonas* sp. DC14 isolated from the Caspian Sea. Int J Environ Sci Technol 12:3895–3904
Mandal D (2006) The use of microorganisms for the formation of metal nanoparticles and their application. Appl Microbiol Biotechnol 69:485–492
Meena KK, Sorty AM, Bitla UM, Choudhary K, Gupta P, Pareek A, Singh DP, Prabha R, Sahu PK, Gupta VK, Singh HB, Krishanani KK, Minhas PS (2017) Abiotic stress responses and microbe-mediated mitigation in plants: the omics strategies. Front Plant Sci 8:172
Mohanpuria P, Rana NK, Yadav SK (2008) Biosynthesis of nanoparticles: technological concepts and future applications. J Nanopart Res 10:507–517
Mukherjee P, Ahmad A, Mandal D, Senapati S, Sainkar SR, Khan MI, Parishcha R, Ajaykumar PV, Alam M, Kumar R (2001) Fungus-mediated synthesis of silver nanoparticles and their immobilization in the mycelial matrix: a novel biological approach to nanoparticle synthesis. Nano Lett 1:515–519
Mukherjee P, Senapati S, Mandal D, Ahmad A, Khan MI, Kumar R, Sastry M (2002) Extracellular synthesis of gold nanoparticles by the fungus *Fusarium oxysporum*. Chem Bio Chem 3:461–463
Muratova AYU, Turkovskaya OV, Antonyuk LP, Makarov OE, Pozdnyakova LI, Ignatov VV (2005) Oil-oxidizing potential of associative rhizobacteria of the genus *azospirillum*. Microbiol 74:210–215
Mustafa S, Kabir S, Shabbir U (2019) Plant growth promoting rhizobacteria in sustainable agriculture: from theoretical to pragmatic approach. Symbiosis 78:115–123
Nacoon S, Jogloy S, Riddech N (2020) Interaction between phosphate solubilizing bacteria and arbuscular mycorrhizal fungi on growth promotion and tuber inulin content of *helianthus tuberosus* L. Sci Rep 10:4916
M. M. Naik , M. S. Prabhu , S. N. Samant , P. M. Naik and S. Shirodkar (2017), *Thalassas: an International Journal of Marine Sciences*, 33 , 73–80.
Nair B, Pradeep T (2002) Coalescence of nanoclusters and formation of submicron crystallites assisted by *Lactobacillus* strains. Cryst Growth Des 2:293–298
Narayanan KB, Sakthivel N (2010) Biological synthesis of metal nanoparticles by microbes. Adv Coll Interface Sci 156:1–13
Nicomrat D, Tharajak J, Kanthang P (2016) Microbial comparison in synthesizing gold nanoparticles for harvesting gold in wastewater system. Appl Mech Mater 848:52–55
Ojuederie OB, Babalola OO (2017) Microbial and plant-assisted bioremediation of heavy metal polluted environments: a review. Int J of Env Res Pub Health 14(12):1504

Olanrewaju OS, Glick BR, Babalola OO (2017) Mechanisms of action of plant growth promoting bacteria. World J Microbiol Biotechnol 33:197

Padrón-Rodríguez L, Arias-Mota RM, Medel-Ortíz R, de la Cruz-Elizondo Y (2020) Interaction with arbuscular mycorrhizal and phosphate solubilizer fungi in Canavalia ensiformis (Fabaceae). Botanical Sci 98:278–287

Panwar M, Tewari R, Nayyar H (2014) Microbial Consortium of Plant Growth-Promoting Rhizobacteria Improves the Performance of Plants Growing in Stressed Soils: An Overview. In: Khan M, Zaidi A, Musarrat J (eds) Phosphate Solubilizing Microorganisms. Springer, Cham

Prasad R, Bhattacharyya A, Nguyen QD (2017) Nanotechnology in sustainable agriculture: recent developments, challenges, and perspectives. Front Microbiol 8:1014

Quan X, Jian YL, Jian-Min I, HuoYan ZHOU, ChangWen WANG, Xiao QC (2008) Enhancement of phosphorus solubility by humic substances in ferrosols. Pedosphere 18:533–538

Raliya R (2013) Rapid, low-cost, and ecofriendly approach her for iron nanoparticle synthesis using *Aspergillus oryzae* TFR9. J Nanoparticles Article ID 141274:1–4

Raliya R, Biswas P, Tarafdar JC (2015) TiO2 nanoparticle biosynthesis and its physiological effect on mung bean (*Vigna radiata* L.) Biotechnol. Rep 5:22–26

Ramadoss D, Lakkineni VK, Bose P, Ali S, Annapurna K (2013) Mitigation of salt stress in wheat seedlings by halotolerant bacteria isolated from saline habitats. Springerplus 2:6

Ramamoorthy V, Viswanathan R, Raguchander T, Prakasanv SR (2001) Induction of systemic resistance by plant growth promoting rhizobacteria in crop plants against pests and diseases. Crop Protect 20:1–11

Rashid S, Charles TC, Glick BR (2012) Isolation and characterization of new plant growth promoting bacterial endophytes. Appl Soil Ecol 61:217–224

Rawat N, Sharma M, Suyal DC, Singh DK, Joshi D, Singh P, Goel R (2019) Psyhcrotolerant Bio-inoculants and Their Co-inoculation to Improve *Cicer arietinum* Growth and Soil Nutrient Status for Sustainable Mountain Agriculture. J Soil Sci Plant Nutr 19:639–647

Raza W, Yang W, Shen QR (2008) *Paenibacillus Polymyxa*: antibiotics, hydrolytic enzymes, and hazard assessment. J Plant Pathol 90:419–430

Richardson AE, Simpson RJ (2011) Soil microorganisms mediating phosphorus availability. Plant Physiol 156:989–996

Rudresh DL, Shivaprakash MK, Prasad RD (2004) Prasad: effect of combined application of *Rhizobium,* phosphate solubilizing bacterium and *Trichoderma* spp. on growth, nutrient uptake and yield of chickpea (*cicer aritenium* L.). Appl Soil Ecol 28:139–146

Mona Saud AL-Ahmadi (February 21st 2019). Pesticides, Anthropogenic Activities, and the Health of Our Environment Safety, Pesticides - Use and Misuse and Their Impact in the Environment, Marcelo Larramendy and Sonia Soloneski, IntechOpen,

Senapati S, Ahmad A, Khan MI, Sastry M, Kumar R (2005) Extracellular biosynthesis of bimetallic Au-Ag alloy nanoparticles. Small 1:517–520

Seshachala U, Tallapragada P (2012) Phosphate solubilizers from the rhizosphere of *Piper nigrum* L. in Karnataka. India Chil J Agric Res 72:397–403

Shao J, Xu Z, Zhang N, Shen Q, Zhang R (2015) Erratum to: contribution of indole-3-acetic acid in the plant growth promotion by the rhizospheric strain *Bacillus amyloliquefaciens*, SQR9. Biol Fertil Soils 51:331–331

Sharma SB, Sayyed RZ, Trivedi MH, Gobi TA (2013) Phosphate solubilizing microbes: sustainable approach for managing phosphorus deficiency in agricultural soils. Springerplus 2:587

Shulse CN, Chovatia M, Agosto C, Wang G, Hamilton M, Deutsch S, Yoshikuni Y, Blow MJ (2019) Engineered root bacteria release plant-available phosphate from phytate. Appl Environ Microbiol 85:01210–01219

Singh A, Prasad SM (2017) Nanotechnology and its role in agroecosystem: a strategic perspective. Int J Environ Sci Technol 14:2277–2300

Singh B, Satyanarayana T (2011) Microbial phytases in phosphorus acquisition and plant growth promotion. Physiol Mol Biol Plants 17:93–103

Souad Z, Ahmed AB, Lakhdar B (2018) Efficacy of Bacillus subtilis and Trichoderma harzianum combination on chickpea Fusarium wilt caused by F. oxysporum f. sp. ciceris. Arch Phytopathol Plant Protect 51:217–226

Srivastava N, Srivastava M, Mishra PK, Ramteke PW (2016) Application of ZnO nanoparticles for improving the thermal and ph stability of crude cellulase obtained from *Aspergillus fumigatus* AA001. Front Microbiol 7:514

Suárez-Moreno ZR, Vinchira-Villarraga DM, Vergara-Morales DI, Castellanos L, Ramos FA, Guarnaccia C, Degrassi G, Venturi V, Moreno-Sarmiento N (2019) Plant-growth promotion and biocontrol properties of three streptomyces spp isolates to control bacterial rice pathogens. Front Microbiol 10:290

Suleman M, Yasmin S, Rasul M, Yahya M, Atta BM, Mirza MS (2018) Phosphate solubilizing bacteria with glucose dehydrogenase gene for phosphorus uptake and beneficial effects on wheat. PloS one 13(9):e0204408

Surange S, Wollum AG, Kumar N (1997) Characterization of *Rhizobium* from root nodules of leguminous trees growing in alkaline soils. Can J Microbiol 43:891–894

Syed A, Ahmad A (2012) Extracellular biosynthesis of platinum nanoparticles using the fungus *Fusarium oxysporum*. Colloids Surf B Biointerfaces 97:27–31

Tahir M, Naeem MA, Shahid M, Khalid U, Farooq ABU, Ahmad N, Waqar A (2020) Inoculation of pqqE gene inhabiting Pantoea and Pseudomonas strains improves the growth and grain yield of wheat with a reduced amount of chemical fertilizer. J Appl Microbiol 129:575–589

Thakkar KN, Mhatre SS, Parikh RY (2010) Biological synthesis of metallic nanoparticles. Nanomed Nanotechnol Biol Med 6:257–262

Thant S, Aung NN (2018) Phosphate solubilization of *Bacillus megaterium* isolated from non-saline soils under salt stressed conditions. J Bacteriol Mycol Open Access 6:335–341

Trolove SN, Hedley MJ, Kirk GJD, Bolan NS, Loganathan P (2003) Progress in selected areas of rhizosphere research on p acquisition. Aust J Soil Res 41:471–499

Ullah A, Mushtaq H, Ali U (2018) Screening, isolation, biochemical and plant growth promoting characterization of endophytic bacteria. Microbiol Curr Res 2:62–68

Venkateswarlu B, Rao AV, Raina P, Ahmad N (1984) Evaluation of phosphorus solubilization by microorganisms isolated from arid soil. J Indian Soc Soil Sci 32:273–277

Vigneshwaran N (2007) Biological synthesis of silver nanoparticles using the fungus *Aspergillus flavus*. Mater Lett 61:1413–1418

Wani PA, Khan MS, Zaidi A (2007) Synergistic effect of the inoculation with nitrogen-fixing and phosphate-solubilizing rhizobacteria on performance of field-grown chickpea. J Plant Nutr Soil Sc 170:283–287

Whitelaw MA (2000) Growth promotion of plants inoculated with phosphate solubilizing fungi. Adv Agron 69:99–151

Zaidi A, Ahmad E, Khan MS (2014) Role of Phosphate-Solubilizing Microbes in the Management of Plant Diseases. In: Khan M, Zaidi A, Musarrat J (eds) Phosphate Solubilizing Microorganisms. Springer, Cham

Zhang F (2016) Biocontrol potential of *Trichoderma harzianum* isolate T-aloe against *Sclerotinia sclerotiorum* in soybean. Plant Physiol Biochem 100:64–74

Zhang Y, Chen FS, Wu XQ, Luan FG, Zhang LP, Fang XM (2018) Isolation and characterization of two phosphate-solubilizing fungi from rhizosphere soil of moso bamboo and their functional capacities when exposed to different phosphorus sources and pH environments. PLoS ONE 13:e0199625

Zhu FG, Qu L, Hong X, Sun X (2011) Isolation and characterization of a phosphate solubilizing halophilic bacterium *Kushneria* sp. Ycwa18 from DaQiao salt Ern on the coast of the yellow sea of china. Evid Based Complement Alternat Med 615032:1–6

SAGA: A Hybrid Technique to handle Imbalance Data in Software Defect Prediction

Ruchika Malhotra
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
ruchikamalhotra@dtu.ac.in

Ritvik Kapoor
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
ritvikush@gmail.com

Paridhi Saxena
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
paridhi.saxena1999@gmail.com

Parth Sharma
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
parthpsharma99@gmail.com

2021 IEEE 11th IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE) | 978-1-6654-0338-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/ISCAIE51753.2021.9431842

Abstract* — *Software defect prediction has been a concurrent topic in software quality-based research. Predictive models that identify defect prone parts of Software can be evolved from defect data and software metrics. Various studies conducted in the past have explored Machine Learning-based approaches for this purpose but the problem of handling imbalanced defect data without compromising on the model's performance remains at large. In this work, we have proposed, compared, and analyzed a hybrid technique, SAGA(SMOTE + AdaSS + Genetic Algorithm), for solving the imbalance problem faced in software defect prediction. SAGA employs ensemble classification based on feature space partitioning in conjunction with the Synthetic Minority Oversampling technique. Various parameters related to feature space partitioning are optimized using the Genetic Algorithm. The values of ROC-AUC, G-mean, Balance, and Accuracy obtained on open-source datasets confirm the effectiveness of the proposed technique.

Keywords*— *software defect prediction, data imbalance, ensemble, feature space partitioning, Genetic Algorithm, Synthetic Minority Oversampling

I. Introduction

Software defect prediction is a central research topic in software engineering. Defects (or bugs) can occur in any phase of the software development life cycle; the early identification of these defects is very important as in the later stages they become costlier to correct and difficult to detect [1][2]. Also, if a defect goes undetected it can cause failures hampering the correct functioning of the software. This crops up the need for accurate software defect prediction techniques to produce high-quality and economical software that ensures customer satisfaction [3]. The ability to learn from previous experiences and predict probable buggy classes/modules are highly sought after by software development companies because they cannot risk customer dissatisfaction, waste effort and damage the code structure by patching bugs.

To classify classes as defective or defect-free, software data is collected during the development process based on different complexity measurement metrics such as Halstead [4], McCabe [5], Object-Oriented (OO) (CK) metrics [6], but the defect data obtained in this fashion has been found out to be imbalanced with lesser number of minority class entries [7,8]. The Machine Learning(ML) models trained from this one-sided distribution of classes perform poorly in classifying minority classes, hampering the overall performance of the model.

We resolve the aforementioned issues, by addressing the following RQs:

RQ1. Can we modify the ADASS ensemble classifier and use it in conjunction with data sampling to develop an imbalance learning technique that works well with the dual-class defect data?

RQ2. On the basis of reliable statistical tests, how does the proposed technique perform in comparison to ensemble classifiers that involve other data sampling strategies?

PROMISE is an open-source repository of software datasets [9], which has been widely used for software quality related research. We have used datasets of thirteen JAVA-based projects obtained from the PROMISE data repository for answering the research questions.

The attempts to address imbalance learning have historically revolved around sampling, algorithmic and ensemble approaches. We have used elements from all three in formulating SAGA (SMOTE + AdaSS + Genetic Algorithm) a hybrid Machine Learning technique for software defect prediction. Synthetic Minority Oversampling(SMOTE) oversampling [10] is one of the most successful data-level methods, employed to handle imbalanced datasets. Lopez-Garcia, Masegosa, Osaba, Onieva and Perallos [11] showed how AdaSS(Adaptive Splitting and Selection) ensembles can be used to handle data imbalance for multiclass classification. We have modified the AdaSS ensemble classifier to make it suitable for working with defect datasets which are mentioned in section III. Using AdaSS ensemble involves partitioning feature space into clusters and giving right weightage to the results of different classifiers making up ensembles training on each cluster, thus training the algorithm entails resolution of a complex optimization problem We have used prime-parents transference Genetic Algorithm (GA) for the optimization problem in a discussion because it is suitable for multi-objective optimization, is robust to local maxima/minima and works well with both discrete and continuous variables as our feature space shows dual characteristics. Before starting training, we have applied CFS (Correlation-based feature selection) as suggested by Malhotra et. al in [7] to reduce training time and improve the accuracy of our classification.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:34:54 UTC from IEEE Xplore. Restrictions apply.

After training, the performance of SAGA is measured using stable result metrics - Area under Curve (AUC), G-mean, balance and accuracy. To foster a perspective for performance comparison we have drawn on four widely used ensemble techniques that involve data level strategies (RUSBoost, EasyEnsemble, BalanceBagging, and AdaBoost) for comparison with our proposed technique with and without SMOTE. The six methods are then ranked by a non-parametric statistical method, Friedman ranking based on their result metrics.

The rest of the paper is arranged in six subsections. Section 2 reviews related work. Section 3 reviews the datasets used. The setup of our approach is explained in Section 4. Section 5 discusses the evaluation measures used in the project. Section 6 explores the results obtained.

II. Background and Related Work

Predicting software defects is a central area of research in Software Engineering. A recent review over 74 SDP by Zhiqiang et al. [12] shows the different approaches adopted and problems faced in SDP. The problem of class imbalance in the dataset has been a recurrent theme across many studies [2,7,8]. Xu et al. [1] have attempted to solve it using kernel PCA and a weighted extreme learning machine, Malhotra and Kamal [8] have used data level oversampling methods, Malhotra and Jain [7] have used boosting based ensemble methods, Sawalqa, Faris, Aljarah, Alnemer and Alhindawi [13] have used a hybrid SMOTE-ensemble technique to name a few. However, we still feel that there is a scope of improvement in providing more balanced and unbiased defect predictions.

The concept of feature space partitioning remains relatively unexplored in SDP. AdaSS algorithm [14] also works the same principle and has never been used for SDP. Garcia, Masegosa, Osaba, Onieva and Perallos [11] have shown how AdaSS ensembles can be used to handle data imbalance and how we can optimize it using Metaheuristic Techniques to give good predictions.

Now, according to Hamdy and El-Laithy [15], we can increase the performance of an algorithm in SDP by using SMOTE and a feature selection technique. Also, it has been successfully applied in conjunction with classifiers to give hybrid algorithms with better performance in the past[13,16].

III. Data Description

There are various software defect prediction dataset sources available for exploration, some of them include AEEEM, NASA, PROMISE, RELINK, SOFTLAB.

This study is based on some of the datasets provided by the PROMISE Data Repository collected by Jureczko and Madeyski [9]. Each project dataset consists of 20 metrics composed of Object-Oriented (OO) metrics, Chidamber-Kemerer (CK) metrics, McCabe's cyclomatic metrics and others. We have used a subset of 13 datasets from the PROMISE repository, on the precedent set by [17]. Table I gives a brief description of the datasets used. From it, we can understand the size, total number of instances, buggy instances and dataset imbalance in each of the 13 datasets.

TABLE I. Datasets Description

S.No.	Project	Total Instances	KLOC	Buggy Instances	Bug%
P1	poi3.0	439	129.25	279	63.55
P2	log4j1.2	205	38.19	185	90.24
P3	xerces1.2	440	159.25	31	7.05
P4	xalan2.4	723	225.09	79	10.93
P5	ant1.7	742	208.65	166	22.37
P6	synapse1.2	255	53.5	86	33.73
P7	lucene2.4	339	102.85	202	59.59
P8	velocity1.6	229	57.01	34	14.85
P9	skarbonka	45	15.03	9	20
P10	nieruchomosci	27	4.75	10	37.04
P11	ivy2.0	352	87.77	40	11.36
P12	jedit4.3	492	202.36	11	2.24
P13	camel1.6	945	113.06	171	18.1

IV. Experimental Setup

In this section, we give a detailed explanation of the different elements that make up the SAGA in a sequential fashion corresponding to the interaction with the data.

First, we discuss feature selection which constitutes the data preprocessing portion, followed by the SMOTE oversampling technique which is used to stabilize the imbalance in the data. We then move on to the Adaptive Splitting and Selection Algorithm which is then followed by explaining the Ensemble Model architecture used.

A. Feature Selection

Each project of PROMISE Data Repository has 20 features; however, it is to be noted that not all features are relevant to the classification problem, and as such these irrelevant and unnecessary features can act as noise which degrades the performance of the classifier. For this purpose, this study uses CFS (Correlation-based feature selection) approach to obtain the subset of features presented in Table II that most affect the target variable [18][19] for imbalanced datasets. CFS approach ensures that for every project, only those features are considered which have a correlation coefficient above a certain threshold, which is |0.2| with respect to this study.

TABLE II. Features Selected

Project	Features Selected
ant1.7	['moa', 'dam', 'avg_cc', 'ce', 'npm', 'loc', 'wmc', 'max_cc', 'rfc']
poi3.0	['dam', 'rfc', 'npm', 'wmc', 'ic', 'cbm']
log4j1.2	['npm', 'ca', 'cbo', 'dam']

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:34:54 UTC from IEEE Xplore. Restrictions apply.

xerces1.2	['wmcs', 'ce', 'moa', 'lcom', 'loc', 'rfc']
xalan2.4	['ce', 'cbo', 'max_cc', 'moa', 'lcom', 'npm', 'wmc', 'loc', 'rfc']
ant1.7	['moa', 'dam', 'avg_cc', 'ce', 'npm', 'loc', 'wmc', 'max_cc', 'rfc']
synapse1.2	['wmc', 'dam', 'moa', 'max_cc', 'cbo', 'ce', 'loc', 'rfc']
lucene2.4	['ca', 'cbm', 'loc', 'ic', 'wmc', 'dam', 'rfc', 'cbo', 'npm', 'ce']
velocity1.6	['npm', 'max_cc', 'moa', 'lcom', 'ce', 'amc', 'wmc', 'rfc', 'loc']
skarbonka	['wmc', 'dit', 'npm', 'lcom', 'dam', 'rfc']
nieruchomosci	['wmc', 'dit', 'lcom', 'cbo', 'ce', 'max_cc', 'moa', 'avg_cc', 'amc', 'rfc', 'loc']
ivy2.0	['moa', 'npm', 'cbo', 'ce', 'max_cc', 'wmc', 'rfc', 'loc']
jedit4.3	['cbm', 'loc', 'wmc', 'npm', 'ca', 'moa', 'cbo', 'rfc', 'ce']
camel1.6	['ca', 'lcom', 'cbo', 'avg_cc', 'noc', 'npm', 'rfc', 'loc', 'wmc', 'max_cc']

B. SMOTE

This is the first constituent that makes up the SAGA. Software defect prediction datasets are small in size and imbalanced in nature. On the basis of these characteristics, there are several oversampling techniques used in software defect prediction including random oversampling, SMOTE, and ADASYN approach. Based on the excellent performance of SMOTE in previous studies we have used it for data modification [15,20].

C. Adaptive Splitting and Selection Algorithm

In this section, we discuss modified adaptive splitting and selection algorithm as used in SAGA. AdaSS employs the fact that the performance of a classifier is inversely proportional to the variance because the training instances if spread over a small range are likely to build more accurate classification rules. These local competencies can be achieved by splitting the feature space into partially overlapping sub-datasets called clusters.

For this, let us assume that the entire feature space, represented by X, can be divided into a set of H clusters, such that,

$$X = \bigcup_{h=1}^{H} \hat{X}_{h,\ \forall a,b\in\{1,\ldots,H\},\ a\neq b,\ \hat{X}_a} \cap \hat{X}_b \neq \phi \quad (1)$$

where $\hat{X}_h$ denotes the h-th cluster. Each cluster is defined using dimension limit matrix dimension C_h. On each of these clusters, we train a local classifier $\underline{\Psi}_h$.

For the application purposes of this work, each classifier $\underline{\Psi}_h$ is an Ensemble Model. We have discussed the composition of ensemble models in further sections. The only external parameter associated with AdaSS is the number of clusters (H). A larger value of H can cause overfitting and unnecessarily increase the training time while taking a lower value of H defeats the whole purpose of clustering. On the pretext of the mentioned facts, the value of H is set to 7.

Moving forward we will understand the classification rules. For each data instance x the classifying ensemble is chosen based on it's Euclidian Distance from centroids of clusters. which are defined as $C_h = \{c_1, \ldots, c_D\}$, here D is the feature space dimension. So,

$$Cno(x) = \arg\ min_{h=1}^{H} \text{distance}\ (x, C_h) \quad (2)$$

here Cno(x) denotes the cluster used for classifying x.

Hence, the classification result of a global combined classifier Ψ, for a data instance x is given by:

$$\Psi(x) = \underline{\Psi}_{Cno(x)}(x) \quad (3)$$

For local ensembles, we already know $\underline{\Psi}$ with a number in subscript ($\underline{\Psi}_h$) is used for referring to the local ensemble classifiers and each of these ensembles is in turn composed of k(base) classifiers represented using Ψ with a superscript (Ψ^1, Ψ^2, ..., Ψ^k). For a given instance x ∈ X, the local classifier is chosen based on cluster centroids. Each local classifier uses a discriminative method $\hat{D}h(x)$ to decide whether x belongs to class i ∈{0,1} by generating a weighted score of k functions $\hat{D}(l)(x)$ corresponding to k base classifiers. The discriminative functions use W, a matrix of k rows and h columns for generating weighted classifications. So,

$$W_h = [w^1{}_h, w^2{}_{h,} \ldots, w^k{}_h(m)] \quad (4)$$

W_h the weighted array for h^{th} cluster. Now,

$$\hat{D}_h(x) = \sum_{l=1}^{k} w^l{}_h(i)\ F^{(l)}(x) / \sum_{l=1}^{k} w^l{}_h(i) \quad (5)$$

We decide value $\underline{\Psi}_h(x)$ on the basis of $\hat{D}_h(x)$ being greater or less than 0.5.

The dataset of N samples is divided into training and testing sets, to find the right balance of bias and variance in our results.

For the global combined classifier optimization criteria, O on the training dataset is given by:

$$O(\Psi) = [W, CC] \quad (6)$$

The classification model is developed by resolving this complex optimization problem.

D. Optimization Variables

Cluster limits matrix CC and averaging weights matrix W are the variables that need to be optimized. CC(H*d*l) defines the data space partitions also called clusters on which local classifiers are trained here we an additional dimension l=2 to account for both the upper and lower limits of the dimension d and W (H x k) contains the weights on the basis of which the contribution of base classifiers is controlled. For generating H data space partitions of dimension d x 2. We use a loop inside which a d x 2-dimensional matrix is constructed with random variables and if more than a threshold (0.6) amount of points x ∈*X* are present in this feature space, it is added to

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:34:54 UTC from IEEE Xplore. Restrictions apply.

matrix CC. W is initialized using a H x k random integers matrix.

E. Ensemble Methods

Ensembles of base classifiers are used corresponding to each cluster. To calculate the prediction each ensemble is using a weighted average of the predictions given by each method.

The classifiers that constitute the ensemble are:

- *SVM*

  Support Vector Machines train towards generating the maximum margin hyperplane using Support Vectors. As we have a high dimensional feature space and the data is not linearly separable, thus, a polynomial-based kernel is used for the SVM.
- *K- Nearest Neighbor*

  This study uses a 3-NN and computes Euclidean distance to compute the nearest neighbor.
- Neural Network

  The size of the input layer equal to the dimensions in the feature space and the size of the output layer equal to 1, considering binary classification. The model uses the He-initialization of parameters which initializes the weights to random scaled values and using L2-regularisation. The model utilizes gradient descent for backpropagation. The number of neurons is decreased gradually through the hidden layers over a certain scale.

F. Genetic Algorithm

The objective of the optimization problem is to learn the best combination of Cluster dimension limits (CC) and ensemble weights (W) that maximizes the values of our result metrics. To solve it, we used Genetic Algorithm [23]. In this Metaheuristic technique, each variable to be optimized is called a chromosome; the set of all chromosomes is called a person and a set of persons is called a population, each new iteration of the population is called a generation. The goal here is to select people in the previous generation called parents by best performance and generate the next progeny generation using the crossover.

Pseudo-Code:

Algorithm-1

Data : POPsize, Tmax

Result : Best Individual found

1. $POP_0 \leftarrow$ Initialize(POP_{size})
2. $t \leftarrow 0$
3. Evaluate(POP_0)
4. while $t < T_{max}$ do
   5. Top_parents $\leftarrow$ selection(POP_t)
   6. Offspring $\leftarrow$ crossover(POP_t)
   7. Offspring $\leftarrow$ mutation(POP_t)
   8. Evaluate(POP_t)
   9. $POP_{t+1} \leftarrow$ UpdatePopulation(POP_t, Offspring, Top_parents)
   10. Add the generated offspring and Top_parents into new generation POP_{t+1}
   11. $t \leftarrow t + 1$

end

Fig. 1: GA Algorithm

The mutation is also allowed to learn new characteristics not present in the parent generation. We have explained the process in Figure 1. We have used a population size of 50 containing variables generalized using the process explained previously. In each generation, we train the modified AdaSS ensemble and the select best-performing parents using Steady-state selection [24] over Type-II error biased fit function for handling imbalance. Out of the best-performing parents randomly 2 are selected for chromosome crossover to produce CC and W metrics belonging to child generation.

Uniform-mutation [25] is applied to these offsprings based on the mutation-probability. Two of the best-performing parents are also propagated to the child generation. This process is repeated over 100 generations and the values corresponding to the best performer in the last generation are selected as final values of cluster dimension limits and ensemble weights for SAGA.

V. Evaluation Measures

Many researchers have corroborated that accuracy can be insufficient and even a misleading performance metric when dealing with imbalanced datasets [21][22]. To handle the imbalance issue in the datasets, we have used AUC, Balance, and G-Mean [8] along with accuracy to assess the performance of the model.

- Sensitivity is the percentage of correctly classified defective instances, $Sensitivity = \frac{TP}{TP + FN}$.
- Specificity is the percentage of correctly classified non-defective instances, $Specificity = \frac{TN}{FP + TN}$.
- G-Mean is the geometric mean of Sensitivity and Specificity, $G - Mean = \sqrt{Sensitivity \times Specificity}$.
- Balance is a modified performance evaluator obtained from FPR (false positive rate) and Precision.

$$FPR = \frac{FP}{FP + TN}$$

$$Precision = \frac{TP + TN}{TP + FP + FN + TN}$$

$$Balance = 1 - \sqrt{\frac{(0 - FPR)^2 + (1 - Precision)^2}{2}}$$

- AUC stands for Area Under ROC (receiver operating characteristic) curve. The ROC curve is a plot of Sensitivity (True Positive Rate) vs (1 - Specificity)

TABLE III. Performance of Various Datasets

ACCURACY	P1	P3	P4	P2	P5	P6	P7	P8	P9	P10	P12	P11	P13
RUSBoost	0.818	0.773	0.735	0.808	0.565	0.719	0.565	0.638	0.667	0.857	0.894	0.807	0.650
EasyEnsemble	0.755	0.664	0.669	0.673	0.559	0.734	0.576	0.672	0.750	0.857	0.683	0.693	0.620

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:34:54 UTC from IEEE Xplore. Restrictions apply.

BalanceBagging	0.764	0.745	0.762	0.731	0.629	0.719	0.576	0.724	0.583	0.714	0.862	0.864	0.734
AdaBoost	0.764	0.800	0.796	0.885	0.704	0.719	0.647	0.690	0.750	0.857	0.951	0.886	0.781
SAGA	0.871	0.839	0.834	0.871	0.729	0.847	0.752	0.855	0.944	0.889	0.921	0.897	0.755
Non-Smote	0.791	0.839	0.839	0.885	0.833	0.844	0.753	0.759	0.917	0.571	0.764	0.909	0.823
ROC-AUC													
RUSBoost	0.815	0.714	0.567	0.601	0.466	0.690	0.561	0.702	0.400	0.900	0.462	0.695	0.495
EasyEnsemble	0.758	0.738	0.703	0.670	0.547	0.740	0.579	0.683	0.850	0.900	0.353	0.708	0.609
BalanceBagging	0.773	0.697	0.666	0.630	0.533	0.709	0.584	0.754	0.550	0.650	0.445	0.727	0.617
AdaBoost	0.752	0.622	0.557	0.500	0.505	0.696	0.638	0.575	0.650	0.900	0.492	0.623	0.523
SAGA	0.888	0.870	0.887	0.899	0.768	0.892	0.805	0.906	0.994	0.972	0.940	0.925	0.801
Non-Smote	0.836	0.871	0.739	0.570	0.672	0.822	0.753	0.671	0.796	0.500	0.885	0.849	0.523
G-MEAN													
RUSBoost	0.815	0.708	0.502	0.538	0.430	0.673	0.560	0.695	0.000	0.894	0.000	0.679	0.430
EasyEnsemble	0.758	0.728	0.701	0.670	0.546	0.739	0.579	0.683	0.837	0.894	0.000	0.708	0.609
BalanceBagging	0.771	0.693	0.649	0.617	0.503	0.706	0.580	0.753	0.548	0.632	0.000	0.704	0.589
AdaBoost	0.749	0.551	0.412	0.000	0.352	0.686	0.632	0.548	0.632	0.894	0.000	0.515	0.326
SAGA	0.856	0.841	0.833	0.869	0.729	0.847	0.752	0.856	0.949	0.913	0.914	0.892	0.755
Non-Smote	0.796	0.841	0.653	0.438	0.605	0.831	0.729	0.661	0.816	0.000	0.873	0.693	0.154
BALANCE													
RUSBoost	0.815	0.714	0.567	0.601	0.466	0.690	0.561	0.702	0.400	0.900	0.462	0.695	0.495
EasyEnsemble	0.758	0.738	0.703	0.670	0.547	0.740	0.579	0.683	0.850	0.900	0.353	0.708	0.609
BalanceBagging	0.773	0.697	0.666	0.630	0.533	0.709	0.584	0.754	0.550	0.650	0.445	0.727	0.617
AdaBoost	0.752	0.622	0.557	0.500	0.505	0.696	0.638	0.575	0.650	0.900	0.492	0.623	0.523
SAGA	0.859	0.842	0.833	0.870	0.728	0.847	0.756	0.858	0.950	0.917	0.917	0.895	0.755
Non-Smote	0.799	0.842	0.685	0.579	0.661	0.832	0.738	0.688	0.833	0.500	0.881	0.731	0.509

VI. Results

We use the values obtained from training the SAGA model across the generations to evaluate its performance in terms of accuracy, ROC-AUC, G-Mean and Balance values of the models and various oversampling techniques are presented in table III. The research questions are answered in the sections below.

A. *RQ1. Can we modify the AdaSS ensemble classifier and use it in conjunction with data sampling to develop an imbalance learning technique that works well with the dual-class defect data?*

While analyzing the results, lesser attention is paid to accuracy values discussed in the previous section. From the results table obtained above, it is to be noted that for SAGA, the range of AUC-ROC is 0.768 to 0.994, the range of G-MEAN is 0.729 to 0.949 and the range of Balance is 0.728 to 0.95 over the chosen datasets. So, it can be concluded that the proposed model has handled the problem well.

B. *RQ2. On the basis of reliable statistical tests, how does the proposed technique perform in comparison to other ensemble classifiers that involve data sampling strategies?*

There are various ensemble methods using sampling techniques available that can be used to handle the data imbalance. Table III shows the results of AdaBoost, BalanceBagging, RSUBoost and Early Ensemble with SAGA that have been used for the comparison. On top of the Friedman test, the ranking method is employed to compare the performance of these techniques.

Table IV shows the Friedman test rankings across the 13 selected promise datasets. The classifiers are ranked based on their performance evaluation using accuracy, ROC-AUC, G-mean, and Balance. The mean rank for each technique and each performance metric is represented. A value of 0.05 is used for α which denotes a Confidence of 95%.

TABLE IV. Friedman Test Results

Rank	Accuracy	ROC-AUC	G-Mean	Balance
Rank 1	SAGA (5.231)	SAGA (5.923)	SAGA (5.923)	SAGA (5.923)
Rank 2	NonSmote (4.923)	Non-Smote (4.077)	Easy Ensemble (3.846)	Non-Smote (3.923)

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:34:54 UTC from IEEE Xplore. Restrictions apply.

Rank 3	ADABoost (4.154)	Easy Ensemble (3.462)	Non-Smote (3.615)	Easy Ensemble (3.462)
Rank 4	Balance Bagging (2.538)	Balance Bagging (3.154)	Balance Bagging (3.154)	Balance Bagging (3.154)
Rank 5	RUSBoost (2.385)	RUSBoost (2.231)	RUS Boost (2.385)	RUS Boost (2.308)
Rank 6	Easy Ensemble (1.769)	ADA Boost (2.154)	ADA Boost (2.077)	ADA Boost (2.231)
p-values	0.00041	0.0127	0.0216	0.0177

- Since, we have obtained p-values less than 0.05 for accuracy, ROC-AUC, G-Mean, and balance null hypothesis considered while doing the Friedman test is refuted. Thus, there is a significant difference between the predictive capabilities of model performances.
- SAGA secured first place in the average ranking by a large margin across ROC-AUC, G-mean, balance and accuracy outperforming all the other techniques.
- We see that without SMOTE other components making SAGA were able to efficiently handle data imbalance in most of the data sets, getting rank 2 behind SAGA in ROC-AUC, balance, accuracy and rank 3 behind Easy Ensemble in G-mean in the Friedman Test. Also, it is evident from table III that in some datasets its performance was not enough alone and the use of SMOTE oversampling has a significant impact on good performance displayed by SAGA.

VII. Conclusion

SAGA has demonstrated a tremendous improvement in all aspects of performance as compared to the methods currently being used. We have reached this conclusion by doing detailed experiments with 13 widely used open-source JAVA datasets. and comparing results using non-parametric Friedman-Test.

We have also analyzed the performance boost provided by using SMOTE and proved even without using it the GA Optimized modified AdaSS ensemble has decent performance for binary classification and imbalance learning.

The set of methods used in the proposed technique work on different difficulties faced with the software dataset, to give good results. ROC, Balance and G-mean scores indicate that we were able to efficiently handle the problem of imbalanced datasets and the accuracy scores prove that we did not overfit the minority classes in doing so.

References

[1] Zhou Xu, Jin Liu, Xiapu Luo, Zijiang Yang, Yifeng Zhang, Peipei Yuan, et al., “Software defect prediction based on kernel PCA and weighted extreme learning machine”, Information and Software Technology, vol.106, pp 182-200, 2019.

[2] H. A. Al-Jamimi, "Toward comprehensible software defect prediction models using fuzzy logic," IEEE International Conference on Software Engineering and Service Science (ICSESS) Beijing, vol.7, pp. 127-130, 2016.

[3] Zhang, Z., Jing, X. & Wang, T., “Label propagation based semi-supervised learning for software defect prediction”, Autom Softw Eng, vol. 24, pp.47–69, 2017.

[4] M. H. Halstead, Elements of Software Science. New York, Elsevier, 1977.

[5] T. J. McCabe, “A complexity measure”, IEEE Trans. Software Eng., vol. SE-2, 1976.

[6] Chidamber S. R. and Kemerer C. F., “A metrics suite for object oriented design”, IEEE Trans. on Software Engineering, vol. 20, no. 6, pp 476–493, 1994.

[7] R. Malhotra and J. Jain, "Handling Imbalanced Data using Ensemble Learning in Software Defect Prediction," 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, pp. 300-304, 2020.

[8] Ruchika Malhotra, Shine Kamal, An empirical study to investigate oversampling methods for improving software defect prediction using imbalanced data, Neurocomputing, vol. 343, pp. 120-140, 2019.

[9] M. Jureczko and L. Madeyski, “Towards identifying software project clusters with regard to defect prediction,” in PROMISE, pp. 1–10, 2010.

[10] N.V. Chawla, K.W. Bowyer, L.O. Hall and W.P. Kegelmeyer, "SMOTE: Synthetic Minority Over-Sampling Technique", J. Artificial Intelligence Research, vol. 16, pp. 321-357, 2002.

[11] P. Lopez-Garcia, A. D. Masegosa, E. Osaba, E. Onieva and A. Perallos, "Ensemble classification for imbalanced data based on feature space partitioning and hybrid metaheuristics", Int. J. Speech Technol., vol. 49, no. 8, pp. 2807-2822, Aug. 2019.

[12] Li, Zhiqiang, Jing, Xiao-Yuan, Zhu, Xiaoke: “Progress on approaches to software defect prediction”, IET Software, vol. 12, no. 3, pp. 161-175, 2018.

[13] H. Al Sawalqa, H. Faris, I. Aljarah, L. Alnemer and N. Alhindawi, "Hybrid SMOTE-Ensemble Approach for Software Defect Prediction", Computer Science On-line Conference, pp. 355-366, April 2017.

[14] K. Jackowski and M. Wozniak, "Algorithm of designing compound recognition system on the basis of combining classifiers with simultaneous splitting feature space into competence areas", Pattern Analysis and Applications, vol. 12, pp. 415-425, 2009.

[15] Abeer Hamdy and Abdulrahman El-Laithy, "Smote and feature selection for more effective bug severity prediction", International Journal of Software Engineering and Knowledge Engineering, vol. 29, no. 6, pp. 897-919, 2019.

[16] C. W. Yohannese, T. Li, M. Simfukwe and F. Khurshid, "Ensembles based combined learning for improved software fault prediction: A comparative study," 2017 12th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), pp. 1-6, 2017.

[17] Ghotra, B., McIntosh, S., Hassan, A.E., “Revisiting the impact of classification techniques on the performance of defect prediction models” IEEE/ACM 37th IEEE International Conference on Software Engineering, vol. 1, pp. 789-800, May 2015.

[18] Yu, Q., Jiang, S., Wang, R., “A feature selection approach based on a similarity measure for software defect prediction” Frontiers Inf Technol Electronic Eng 18, pp. 1744–1753, Nov 2017.

[19] Lei Yu and Huan Liu, “Feature Selection for High-Dimensional Data A Fast Correlation-Based Filter Solution”, Machine Learning Proceedings of the Twentieth International Conference, August 21-24, 2003.

[20] Hamad Alsawalqah, Hossam Faris, Ibrahim Aljarah, Loai Alnemer, and Nouh Alhindawi, “Hybrid SMOTE-Ensemble Approach for Software Defect Prediction”, Software Engineering Trends and Techniques in Intelligent Systems, Proceedings of the 6th Computer Science On-line Conference, Vol. 3, pp. 355-366, 2017.

[21] L. Chen, B. Fang, Z. Shang and Y. Tang, "Tackling class overlap and imbalance problems in software defect prediction", Softw. Quality J., vol. 26, no. 1, pp. 97-125, 2018.

[22] Q. Song, Y. Guo and M. Shepperd, "A Comprehensive Investigation of the Role of Imbalanced Learning for Software Defect Prediction," in IEEE Transactions on Software Engineering, vol. 45, no. 12, pp. 1253-1269, 2019.

[23] David E. Goldberg, “Genetic Algorithms in Search Optimization and Machine Learning”, Addison Wesley, pp. 41, 1989.

[24] Durillo, J.J., Nebro, A.J., Luna, F. and Alba, E., “On the effect of the steady-state selection scheme in multi-objective genetic algorithms”, International Conference on Evolutionary Multi-Criterion Optimization, pp. 183-197, April 2009..

[25] Helmuth, T., McPhee, N.F. and Spector, L., “Program synthesis using uniform mutation by addition and deletion”, Proceedings of the Genetic and Evolutionary Computation Conference, pp. 1127-1134, July 2018.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:34:54 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021)
IEEE Xplore Part Number: CFP21K74-ART; ISBN: 978-0-7381-1327-2

Sequential Recommendation with Temporal Context via Convolutional Sequence Embedding

Dr. Aruna Bhat
Computer Engineering
Delhi Technological University
Delhi, India
aruna.bhat@dtu.ac.in

Rishabh Chandra
Computer Engineering
Delhi Technological University
Delhi, India
rishabhchandra_2k17co264@dtu.ac.in

Kanika
Computer Engineering
Delhi Technological University
Delhi, India
kanika_2k17co151@dtu.ac.in

2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) | 978-1-6654-1272-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICCS51141.2021.9432323

***Abstract*— Human behavior consists to a large extent of repeated temporal patterns. What a person is going to do next often depends on what time it is as well as what things she/he has interacted with in recent past. Often times, in many fields it is required to predict what actions or items a user is most likely to engage with in near future. To solve this problem, there are many deep learning approaches such as RNN, LSTM [2, 3] but they are usually computationally more demanding as well as difficult to train. Other sequential recommendation methodologies like CASER [1] does not incorporate temporal context. To overcome these limitations, this paper has proposed a convolutional neural network based deep learning approach to solve this problem. The experiments on public dataset demonstrated that our model which takes in account the temporal context performs better than CASER [1] without temporal context. Also, this research work has demonstrated the application of the proposed model as an on-device solution to enable voice assistants or other host applications to proactively provide recommendations and suggestions based on users' past activity, routine and current timestamp.**

Keywords*— *Recommender System; Sequential Prediction; Convolutional Neural Networks; Feed Forward Neural Network; Deep Learning; Embedding Lookup; Temporal Data Feature Engineering; On-Device Intelligence.

I. INTRODUCTION

In 2020, the number of mobile users worldwide stood at 6.95 billion, with forecasts suggesting this is likely to rise to 7.1 billion by 2021. In 2024, the number of mobile users worldwide is projected to reach 7.41 billion [14]. On our mobile devices, this research work interacts with plethora of applications and activities. Each category of app can be seen as an intent. For example, checking Facebook or LinkedIn can be put under social media intents, similarly using Spotify or streaming Netflix can be labelled as entertainment-based intents. Software modules like voice assistants, in-app recommendations etc., often provide suggestions on what the user want to do next based on their past interactions and other contextual understandings.

Suppose user has a routine of checking the weather forecast for the day when she wakes up. If an algorithm is able to learn about the user pattern and predict that at this moment user is most likely to engage in a particular intent then based on her preferences, the relevant data can be fetched from content providers and provide it to the user, so that when she wakes up, she can directly see the updates. This idea can be extended to many other use cases and provide suggestions in variety of domains ranging from booking cabs for you for office in morning to recommending good places to have dinner on Saturday evening.

Coming up with a learning algorithm which can predict what actions or items a user is most likely to engage with in near future depends on user activity sequence as well as temporal context of events. It is very important to learn sequence of activities because most of the time, there is some proclivity in user's behavior. For example, a user might like to see weather forecast after reading news or might look for places after gym session. To learn patterns, there are many deep learning as well as machine learning approaches. But most of these model does not use temporal information for making predictions. But if analyzed, it makes so much sense that a person's day to day activity is so much dependent on time. For example, a person is more likely to go to bar on Friday night than Monday morning and it is definitely less likely that a person will go to office on Sunday evening

There are many deep learning approaches that aims at learning the sequences for prediction, but they do not consider temporal context. This paper proposes a model, inspired from a state-of-the art model called CASER [1], which incorporates both sequential history as well as temporal context for predicting top N intents that a user is most likely to engage with in near future. With correct prediction of intents, it can be handled in a way that it can interact with user in a proactive manner.

II. BACKGROUND AND RELATED WORK

Creating a recommender system has always been an easy job, but considering sequential patterns, time of actions, location of actions is regarded as a challenging task. The earliest approach is Markov Chain based model [4, 5, 8] where from N previous actions Markov Chain recommends actions to the user. The probability of an actions depends only on the previous action. Markov Chain based model and its variant [4, 5, 8] has few disadvantages. One, concept of temporal context is missing. It doesn't hold time as a parameter of deciding the next action. Second, this model focuses on point-level sequence only. It doesn't take into account the effect of combination of previous actions. For example, there is a higher probability that the user will buy eggs if he has bought milk

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:36:09 UTC from IEEE Xplore. Restrictions apply.

and flour together than buying flour or milk separately. Factorized Sequential Prediction with Item Similarity Model (Fossil) [6] uses weighted sum aggregation over previous actions to predict the next possible action.

Also, top-N recommendation [12], matrix factorization [11, 13], collaborative filtering [9] , doesn't model the order of events hence that too doesn't consider the sequence of actions. For session-based recommendation Recurrent Neural Networks (RNN) [2, 3] is used. In RNN based models, previous actions have some kind of relation among them. It works well when the model has considerable sequential patterns. None of these considers the concept of temporal recommendation. Also, RNN based models are difficult to train and often are computationally heavy so they cannot be deployed as on device solutions. Apart from this there are many earlier works done on sequential pattern mining [7, 10].

One recent work done is state-of-the-art model CASER [1]. This model is very successful in capturing union-level as well as point-level sequential patterns. But it doesn't take in account the temporal context and the effect of the time of recommendation and the time of the occurrence of the events. The next section has proposed a methodology, which is inspired from CASER [1] and incorporates temporal context also.

III. PROPOSED METHADOLOGY

The proposed model uses Convolutional Neural Network for learning sequential features in a way that it captures the patterns at both union-level and point-level. Our model's architecture is inspired from state-of-the-art model CASER [1]. The original CASER [1] model ignores the temporal context, but our model architecture takes care of the temporal data also so as to understand the temporal context of the situation. Our model has three main components, Embedding Lookup, Convolutional Layers and Fully Connected Layers. These are explained in subsequent sections.

A. Handling Temporal Data

Temporal data represents a state in time for example when was some particular intent was executed. This data represents time in any dimension for example hour or minute of the day or day of the week or month etc. Temporal data is different from historical data in sense that historical data represents what intents were executed in past and in what sequence. One of the important parts of our methodology is to incorporate temporal context in our model's architecture. Time is inherently cyclical in nature. Hours, minutes, seconds, days of the week all follow cycles. Feeding raw timestamps to our model would not make much sense because 2300 hours and 100 hours though look 22 hours apart but they are in sense only 2 hours apart. It conveys the cyclical nature of the time to our deep learning model. To capture the cycles in time, each temporal feature is converted to create two new features, deriving a sine transformation and a cosine transformation. We need two transformation because if we just use sine transformation, we get this weird side-effect of symmetry. In figure 1, you can see that from sine transformation feature alone, it appears that every horizontal line that we draw across the graph touches two points. So that's why we need two dimensions for a cyclical feature, so cosine comes to rescue. With cosine transformation also, we can break this symmetry [15]. Using these two features together we can capture all the times distinguished from each other.

From the above figures, we can conclude that for each temporal feature we can convert it into two dimensions. In our problem, we consider only two aspects of time, minute of the day and day of the week. So, we represent each temporal context as a vector of 4 values, 2 values representing minute of the day and 2 values representing day of the week.

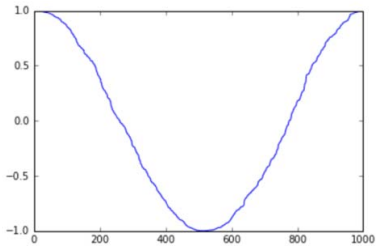


Figure 1: Sine transformation applied on temporal data.

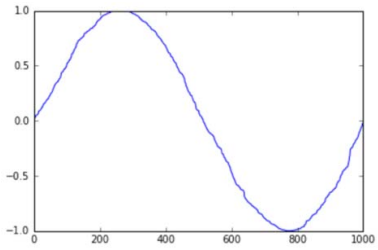


Figure 2: Cosine transformation applied on temporal data.

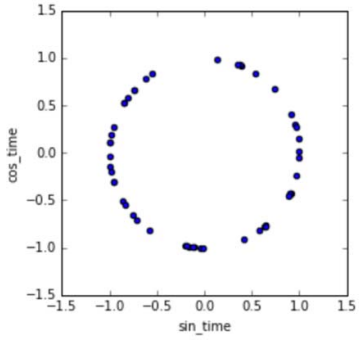


Figure 3: Sine and Cosine transformation both applied on temporal data.

B. Embedding Lookup

Our model uses the embeddings of previous L intents to capture the sequential patterns in latent space. With the help of embedding look-up, it transforms the previous L intents into a d dimension vector space and stack them together resulting in an L x d matrix. Each intent also has its temporal information associated with it. For example, day of the week and time of the day at which that particular intent was interacted with. We transform this temporal information in a vector of 4 dimensions

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:36:09 UTC from IEEE Xplore. Restrictions apply.

(2 for time of the day and 2 for day of the week) so as to capture the cyclical nature of the time as explained before.

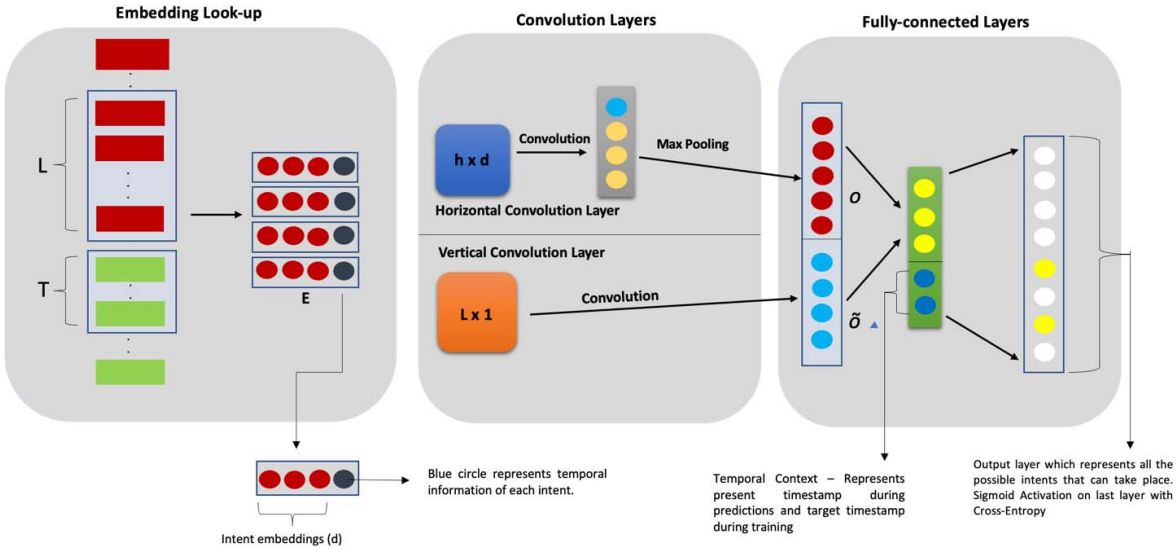


Figure 4: Architecture of our model.

We concat this temporal vector with the embedding of that particular intent which finally results in a matrix E of size L x (d + 4). The embeddings for all intents are learnt simultaneously with all the other weights in the model.

C. Convolutional Layer

In recent times, we have witnessed how convolutional filters of CNN have proved to be a monumental technique in capturing local features for many purposes like image recognition and NLP tasks. Taking inspiration from the same, we treat our L x (d + 4) matrix as an image E of last L intents in the latent space. In this analogy the local features in our image E is nothing but sequential patterns hidden inside it. With the help of horizontal and vertical convolutional filters we can extract the sequential patterns hidden inside it along with the temporal highlights of those patterns.

Horizontal Convolutional Layer. This layer is shown in the upper part of the second component in figure 4. In our architecture we have n filters, each of size h x (d + 4), where h is the height of the filter and can vary from 1 to L. So, if L is 5 and we have 2 filters for each height h, then we will have total of 10 filters. Each of this filter is slide from top to bottom on image E and performs convolution operation to obtain vector c^i for i^{th} filter. After convolution, we apply max pooling operation to extract the maximum value from many values produced by this particular filter. Horizontal convolution layers are responsible for capturing union level sequential patterns of multiple union sizes. Therefore, for all n filters we obtain o as –

$$o = \{\max(c^1), \max(c^2), \dots, \max(c^n)\} \qquad (1)$$

Here max(A) is the max-pooling function applied over vector A.

Vertical Convolutional Layer. This layer is shown in the bottom part of the second component in figure 4. In this layer, we have ñ filters where each filter is of size L x 1. Each filter slides on image E from left to right d + 4 times and all the convolutions by each i^{th} filter is aggregated to produce $ć^i$. Therefore, with vertical filters we aggregated the embeddings of last L intents and they are responsible for capturing point level sequential patterns. In the end, for all ñ filters, we obtain õ as –

$$\tilde{o} = \{ć^1, ć^2, \dots, ć^{\tilde{n}}\} \qquad (2)$$

D. Fully Connected Layer

We combine the output of horizontal convolution layer o and vertical convolutional layers $\tilde{o}$ and feed this information to a fully connected neural network to get more abstract and high-level features. This neural network uses the outputs of convolutional layers as input layer. It has one hidden layer and in this hidden layer we concatenate the timestamp vector. This timestamp vector represents timestamp of the target event during training and during testing it represents the time at which we want our prediction. The last layer is output layer which represents all the possible intents that can take place. Our problem can be classified under multi-label multi-class classification, so we will use sigmoid activation on last layer with cross-entropy as our loss function.

E. Network Training

We are predicting what will happen next based on intents that were interacted in near past and present time context. Each intent is given a unique numeric ID and the time of its

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:36:09 UTC from IEEE Xplore. Restrictions apply.

execution is represented as explained in *Handling Temporal Data* of this section. To train our model, we first take *L* successive intents that the user has interacted with most recently and all the intents that happened in the next *H* hours as target intents, *T*. This is done by a sliding window, which has *L* fixed intents and *T* variable intents (depending on *H*) and each window generates a training instance. We decided to create a window of three hours. Now our target will not be a single intent rather all the intents that happened in next three hours in one hot encoded form. So, in order to train our model for a single data point representing a given timestamp, we take last *L* intents along with their timestamps as input and take all those intents that took place in window of next three hours as target intents.

F. Predictions

Once our model is trained enough, we can start making predictions. To predict what intents are more likely to be interacted in near future, we can take last *L* intents (along with their timestamps) that were interacted in past and current time and feed it to our trained model. Since we are using sigmoid activation in last layer, it will give us probabilities of each possible intent that can happen in next three hours. From this we can take top N intents which have highest probability.

IV. EXPERIMENTS

We did some experiments to verify how well our methodology is performing in finding out future possible intents based on sequential and temporal context.

A. Dataset

Sequential recommendation only matters when data set has some sequential patterns hidden within it. For our experiments we needed a dataset in which we can find some pattern and a routine in user's behavior.

For our experiments we have taken subsets of two publicly available data sets –

1. Foursquare - This dataset contains spatial-temporal information about check-ins in NYC collected for about 10 months. It contains 227,428 check-ins in New York city spanning over 1011 users and 412 venue categories.
2. Gowalla – Similar to Foursquare dataset containing 7000 check-ins spanning over 23 users and 502 item categories.

Categorization or distinguishing objects based on its similar attributes is a type of supervised-ML which is widely used in various fields such as science, medicine, business, and industry to support decision-making [16]. The above two datasets contain spatial-temporal data about different individuals. There is always a pattern or tendency hidden in a person's behavior. For example, a particular person might like to go to gym in morning or some fancy restaurant on Saturday evening. That's why we have taken these two datasets because they have data about where people are going and at what time. We cleaned the data and categorized locations like gym, restaurant etc. as intents.

For our experiments, the first 70% of each user's sequence is used for training, 10% is used for validation and last 20% is used as testing set for evaluating our model's performance on different evaluation metrics.

B. Results

In order to prove if adding temporal context to sequential recommendation helps, we need to compare our model with the original CASER model which doesn't take in account the temporal context. Since our model's architecture is a modification of CASER, it makes more sense to replicate similar experiments so as to compare it with the original CASER model. We use Precision@N and Recall@N to evaluate our model's performance and compare it with CASER. Given a list of top N predictions of future intents for a user, denoted $R_{1:N}$, and the last 20% of actions in her/his sequence (i.e., denoted R (i.e., the test set), Precision@N and Recall@N are computed by

$$Precision@N = \frac{|R \cap R_{1:N}|}{N} \quad (3)$$

$$Recall@N = \frac{|R \cap R_{1:N}|}{|R|} \quad (4)$$

We report the average of these values of all users. $N \in \{1, 3, 5\}$.

To verify if adding temporal context helps in providing better recommendations, we did multiple experiments with same dataset (described in previous section) using these two models –

- Without Temporal Context – The original CASER [1] model which does not account temporal context of the intents.
- With Temporal Context – Our model which incorporates temporal context of the intents for making future recommendations.

Results from our experiments are reported in Table 1 and Table 2.

Table 1: Performance Comparison on Foursquare dataset.

Metric	Without Temporal Context	With Temporal Context	Improv.
Prec@1	0.1796	0.1921	6.9%
Prec@3	0.0823	0.0857	4.1%
Prec@5	0.0552	0.0573	3.8%
Recall@1	0.0757	0.0803	6.1%

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:36:09 UTC from IEEE Xplore. Restrictions apply.

Recall@3	0.1324	0.1370	3.5%
Recall@5	0.1639	0.1702	3.8%

Table 2: Performance Comparison on Gowalla dataset.

Metric	Without Temporal Context	With Temporal Context	Improv.
Prec@1	0.3307	0.4020	21.5%
Prec@3	0.1817	0.2104	15.8%
Prec@5	0.1391	0.1531	10.1%
Recall@1	0.0523	0.0624	19.3%
Recall@3	0.1377	0.1531	11.2%
Recall@5	0.2054	0.2252	9.6%

As we can see from the above two tables, we can clearly see that adding temporal context to the original CASER improves the metrics. Prec@1 is increased by 6.9% in Foursquare's dataset and 21.5% in Gowalla's dataset. This proves that feeding temporal information to our model helps it in making much better predictions as compared to results with no temporal context at all.

V. APPLICATIONS

From above experiments, we can conclude that our methodology can provide predictions on basis of sequential history and temporal context. We can apply this methodology to enable voice assistants or other host applications to proactively provide recommendations and suggestions based on user's past activity, routine and current timestamp.

Suppose user has a routine of checking the weather forecast for the day when she wakes up or maybe she checks for weather forecast every morning between 7am to 9am or maybe after reading news. On weekends, she might usually check AQI in the afternoon. Our model can learn this pattern in user activity. Once we know the user's possible future intents, we can fetch the relevant data from content providers and provide it to the user, so that when she wakes up, she can directly see the updates.

Our model is capable of learning the user activity sequence as well as temporal context. To learn patterns, there are many deep learning approaches such as RNN, LSTM but they are usually computationally more demanding as well as difficult to train.

We did few more experiments in which we applied our methodology to provide proactive suggestions i.e., what information or activity user is most likely to interact with based on user's past activity on device and context understanding.

A. Dataset Curation

We needed a dataset which contains a log of user's interaction with different category of apps (intents) with proper timestamp. We were unable to find any such publicly available dataset due to privacy reasons. We generated our own dataset by describing intent descriptions as input to our generator script. Each intent description has following attributes –

1. Intent name
2. Probability of event happening on different days of week
3. Time range in which that event will happen

```
{
        'intent_name' : 'read_news',
        'probability' : [0.8,0.8,0.8,0.8,0.8,0.2,0.2],
        'start_hour' : '8:00',
        'end_hour' : '10:00'
}
```

Figure 5: Intent description for read_news event that will randomly happen between 8:00AM to 10AM with 80% probability on weekdays and 20% probability on weekends.

Similar to *read_news* on previous slide we make a list of many intent descriptions with some kind of pattern that a normal working person has in his life. Like waking up, reading mail, going to work, watching movies on Friday nights and so on. Obviously, a person's daily activities are not that regular or fixed, so we add noise in our data with some probability. This noise is added using an even description for each intent with a fixed probability.

	intent	timestamp
0	read_news	2020-05-01 07:01:00
1	weather.check	2020-05-01 07:48:00
2	fetch_mails	2020-05-01 10:04:00
3	calendar_events	2020-05-01 10:33:00
4	stock_info	2020-05-01 11:36:00
5	music	2020-05-01 12:39:00
6	music	2020-05-01 16:04:00
7	book_ride	2020-05-01 17:55:00
8	find_place	2020-05-01 18:56:00
9	streaming_service	2020-05-01 21:29:00
10	read_news	2020-05-01 21:57:00
11	weather.check	2020-05-01 22:24:00
12	clock.set_alarm	2020-05-02 05:37:00
13	read_news	2020-05-02 09:15:00
14	fetch_mails	2020-05-02 12:07:00
15	weather.air_quality	2020-05-02 12:36:00
16	streaming_service	2020-05-02 15:43:00

Figure 6: Sample data with 14 unique intents.

B. Evaluation Metrics

To evaluate performance of our model, we need a new evaluation metric for these experiments. In general, most recommendation systems use Precision@N and Recall@N as evaluation metrics. For our purpose, we need something

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:36:09 UTC from IEEE Xplore. Restrictions apply.

similar to Precision@N. Our model will predict top N intents and we will take next three hours' intents as ground truth.

Score@1 is probability of our top 1 predicted intent happening in next three hours and Score@3 is telling us probability of any one of top 3 predicted intents happening in next three hours. Remember that we have also trained our model with next three-hour intents as targets, so that's why we are using next three-hour intents as ground truth for testing too.

C. Results

After training my model with 30 days user's intent data and testing it on 15 days of data, we got following results.

Table 3: Performance of our model on generated dataset.

Metric	Value
Score@1	0.5437
Score@3	0.8621

This means that almost 55% time our model predicts intent that will take place in next three hours and almost 86% time our model predicts intent from which at least one will take place in next three hours. Scores varies with noise and sequential intensity of data as well, but variance is not so huge. With dataset having higher sequential intensity and less noisy data, model can reach much better score.

```
predict_now('./input/sample2.2_MON17',item_map,all_item_count,[[17,1]])
print("Intent History Input -\n")
print(pd.read_csv('./input/sample2.2_MON17'))
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Prediction for 5PM , Monday
1 : music
2 : book_ride
3 : stock_info
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Intent History Input -

                intent           timestamp
0           read_news 2020-06-22 09:02:00
1         fetch_mails 2020-06-22 09:58:00
2     calender_events 2020-06-22 10:17:00
3           book_ride 2020-06-22 10:53:00
4          stock_info 2020-06-22 12:18:00
5  weather.air_quality 2020-06-22 14:02:00
6         fetch_mails 2020-06-22 15:52:00
7   streaming_service 2020-06-22 15:58:00
8               music 2020-06-22 16:15:00
```

```
predict_now('./input/sample2.1_FRI10',item_map,all_item_count,[[10,5]])
print("Intent History Input -\n")
print(pd.read_csv('./input/sample2.1_FRI10'))
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Prediction for 10AM , Friday
1 : fetch_mails
2 : weather.air_quality
3 : stock_info
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Intent History Input -

                intent           timestamp
0          stock_info 2020-06-11 11:39:00
1         fetch_mails 2020-06-11 14:35:00
2           book_ride 2020-06-11 18:13:00
3   streaming_service 2020-06-11 20:53:00
4           read_news 2020-06-11 21:23:00
5     clock.set_alarm 2020-06-11 22:21:00
6  clock.disable_alarm 2020-06-11 23:29:00
7           read_news 2020-06-12 06:15:00
8         fetch_mails 2020-06-12 09:31:00
9     calender_events 2020-06-12 09:59:00
```

Figure 7: For a given intent sequence with timestamp, our model will predict top three intents that are most likely to happen in next three hours.

D. Computation Complexity

On-device machine learning allows you to train as well as run inference directly on the devices, with the benefits of data privacy, regardless of connectivity. History of user activity is very private to any user and no one would like to send such kind of data to any cloud platform. Many intrusion detection systems have been established to detect the network attacks surroundings. There should be an essential of high accuracy to detect those invaders for cyber security [17]. So, to ensure end-users of their data privacy we need a light weight on-device solution which can train as well as run inference directly on user's device.

Successful deep learning models often require significant amounts of computational resources, memory and power to train and run, which presents an obstacle if you want them to perform well on mobile devices. One such factor is often memory requirement of the model for its training as well as inference. Mobile devices have less memory as compared to traditional machines. In order to deploy our model as an on-device solution, we need to estimate memory required especially during training of the model.

Table 3: Model Complexity - Latent Dimensions (LD), Horizontal Filters (H), Vertical Filters(V), Temporal Dimensions (TD)

Hyperparameter Setting	Number of Trainable Parameters
8LD 8H 4V 0TD	4786
8LD 8H 4V 2TD	5758
8LD 8H 4V 4TD	6730
16LD 16H 8V 7TD	26568

Model with best results has 5758 trainable parameters. If we have n trainable parameters and memory needed for storing 32-bit datatype (single-precision floating point) is 4 Bytes, then total memory needed for n parameters will be 4n Bytes for loading 1 input at a time. Using mini batch for training, number of inputs in each batch size, b, memory

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:36:09 UTC from IEEE Xplore. Restrictions apply.

needed for loading b inputs at a time will be 4nb Bytes. A backward pass requires about triple this, because of the need to store:

- the activations and associated gradients for each neuron - these are of equal size.
- the gradients of the weights (parameters) which are the same size as the parameters.
- the value of the momentum, if you're using it.
- some kind of miscellaneous memory

So, final memory requirement will be 12nb Bytes. For 5758 parameters and 64 as size of batch size, we need around 4.42 MB of RAM. From above calculations we can see that our model is light enough to be deployed as on-device solution.

VI. Conclusion and Future Scope

In this paper, we have proposed a deep learning model that captures many important features of sequential recommendation i.e., union-level and point-level sequential trends and also incorporates temporal contextual understanding. There are many deep learning approaches based on RNN, LSTM etc., but they are usually computationally more demanding as well as difficult to train. Other sequential recommendation methodologies like CASER does not incorporate temporal context. These limitations are overcome in our methodology. Our model provides users with relevant suggestions based on temporal contextual understanding as well as on items or activities user engaged most recently with. We also demonstrated application of our algorithm as an on-device solution for enabling mobile devices to provide suggestions in a proactive manner based on user activity.



Figure 8: Flowchart of module for handling weather intent for more interactive suggestions.

In future, we could extend this work by adding spatial context to our model architecture. In our application we could also make modules for handling each specific intent separately to provide more interactive suggestions. For example, figure 8 for handling weather intents. Some ideas about handling other intents for example, asking user to book rides for office etc. We can extend this work by building modules for handling different intents to build a recommendation system which can provide users suggestions and recommendations in a proactive manner.

Acknowledgment

We would like to thank Dr. Aruna Bhat for her able guidance and support throughout the project.

References

[1] Jiaxi Tang and Ke Wang. 2018. Personalized Top-N Sequential Recommendation via Convolutional Sequence Embedding. In *Proceedings of Eleventh ACM International Conference on Web Search and Data Mining*, February 5-9, 2018, Marina Del Ray, CA, USA, (WSDM 2018), 9 pages.
[2] Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. 2015. Session-based recommendations with recurrent neural networks. arXiv preprint arXiv:1511.06939 (2015).
[3] Dietmar Jannach and Malte Ludewig. 2017. When Recurrent Neural Networks meet the Neighborhood for Session-Based Recommendation. *In ACM Conference on Recommender systems*. ACM, 306–310.
[4] R. He and J. McAuley. 2016. Fusing Similarity Models with Markov Chains for Sparse Sequential Recommendation. In *International Conference on Data Mining*. IEEE
[5] Chen Cheng, Haiqin Yang, Michael R Lyu, and Irwin King. 2013. Where You Like to Go Next: Successive Point-of-Interest Recommendation. *In International Joint Conference on Artifcial Intelligence*. 2605–2611.
[6] Lei Zheng, Vahid Noroozi, and Philip S. Yu. 2017. Joint Deep Modeling of Users and Items Using Reviews for Recommendation. *In International Conference on Web Search and Data Mining*. ACM, 425–434.
[7] Rakesh Agrawal and Ramakrishnan Srikant. 1995. Mining sequential patterns. *In International Conference on Data Engineering*. IEEE, 3–14
[8] Steffen Rendle, Christoph Freudenthaler, and Lars Schmidt-Thieme. 2010. Factorizing personalized markov chains for next-basket recommendation. In *International Conference on World Wide Web*. ACM, 811–820
[9] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl. 2001. Item-based collaborative filtering recommendation algorithms. In *International Conference on World Wide Web*. ACM, 285–295.
[10] Jiawei Han, Jian Pei, and Micheline Kamber. 2011. Data mining: concepts and techniques. Elsevier
[11] Andrej Karpathy, George Toderici, Sanketh Shetty, Thomas Leung, Rahul Sukthankar, and Li Fei-Fei. 2014. Large-scale video classification with convolutional neural networks. In *IEEE conference on Computer Vision and Pattern Recognition*. 1725–1732.
[12] Rong Pan, Yunhong Zhou, Bin Cao, Nathan N Liu, Rajan Lukose, Martin Scholz, and Qiang Yang. 2008. One-class collaborative filtering. In *International Conference on Data Mining*. IEEE, 502–511
[13] Ruslan Salakhutdinov and Andriy Mnih. 2007. Probabilistic Matrix Factorization. In *Advances in Neural Information Processing Systems*, *Vol*. 1. 2–1.
[14] Forecast number of mobile users worldwide 2020-2024. Published by S. O'Dea, Sep 22, 2020.
[15] Ian London's Blog: Ian London (2016, June 31). Encoding cyclical continuous feature-https://ianlondon.github.io/blog/encoding-cyclical-features-24hour-time/
[16] Manoharan, Samuel. "Population Based Meta Heuristics Algorithm for Improvement of Feed Forward Neural Network." Journal of Soft Computing Paradigm (JSCP) 2, no. 01 (2020): 36-46.
[17] Many intrusion detection systems have been established to detect the network attacks surroundings [3]. There must be an essential of high accuracy to detect those invaders for cyber security.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:36:09 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021)
IEEE Xplore Part Number: CFP21K74-ART; ISBN: 978-0-7381-1327-2

SIGN LANGUAGE ALPHABET RECOGNITION USING CONVOLUTION NEURAL NETWORK

Mayand Kumar
Department of Software Engineering
Delhi Technological University
Delhi,India
mayandkumar10@gmail.com

Piyush Gupta
Department of Software Engineering
Delhi Technological University
Delhi, India
piyush.gupta1811@gmail.com

Rahul Kumar Jha
Department of CSE
Delhi Technological University
Delhi, India
Heavenboyrj.11@gmail.com

Aman Bhatia
Department of Software Engineering
Delhi Technological University
Delhi, India
aman.b25@hotmail.com

Khushi Jha
Department of CSE
Delhi Technological University
Delhi, India
khushi.ktm2@gmail.com

Bickey Kumar Shah
Department of CSE
Delhi Technological University
Delhi, India
Bickeyshah721@gmail.com

2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) | 978-1-6654-1272-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICCS51141.2021.9432296

Abstract— **Sign Language plays an indispensable role in the lives of people having speaking and hearing disabilities. Recognition of American Sign Language using Computer Vision is very challenging due to its increasing complexity and high intraclass variations. In this paper, convolutional neural networks (CNNs) are used to recognize the ASL Alphabets. This algorithm is useful to recognize it as a deep network, which is expected for the ASL alphabet classification task. Pre-Processing steps of the MNIST dataset are done in the first phase. After the first phase, different important features of pre-processed hand gesture image are computed. In the final phase, depending on the properties computed or calculated in the initial phases, the accuracy and AUC score of the network model with which it can recognize the sign language Alphabets were detected. The proposed CNN network has achieved an AUC score of 0.9981 and an accuracy of 0.9963.**

Keywords— Convolution neural network (CNN), American Sign Language (ASL), ASL number, Data augmentation, Sign language recognition component.

I. INTRODUCTION

The Sign Language Alphabets are created by using hand and facial gestures in order to express the thoughts of speaking and hearing disabled population to standard people, who may not purely understand the sign language. This develops a vast communication gap between normal people and the people with disabilities. There are certain technological developments to help the speaking and hearing disabled people as it is impossible for them to seek help all the time and assist them in their daily activities as a translator. A novel approach can clarify sign gestures into decipherable text with the help of technology so that it becomes understandable for the normal public. This approach will surely reduce the communication gap and initiate the communication between impaired and standard people in a more feasible way. More than 100 different sign languages are utilized across the world for understanding the deaf and speech impaired groups, they are Indian Sign Language, American Sign Language (ASL), Italian Sign Language, and many others. More than thousands and millions of people and about millions of people in India are using Sign Language as their primary mode of communication.

American Sign Language is the widely used sign language and it ranks fourth most used language in North America. It is a surprise to know that not only in the USA, ASL is also used in more than 30 other nations, where English is considered as the primary mode of communication. Around a million people in USA and various parts of the globe use ASL as their primary mode of communication [1]. ASL is quite vast and complex language, which is created by using hand, finger actions, and facial gestures to express the thoughts of speech-impaired people. ASL Language is recognized as an appropriate and original language, where it contains more number of variations like any other language. For example: Italian and German. Sign language is a magnificent piece of interaction art that is favourable to a wide community of speaking and hearing disabled people. The existence of ASL promotes the happiness and hope amongst impaired people and its wide impact is marvellous and eye-catching [2].

ASL is a set of 26 gesture signs popularly known as American Manual Alphabet, which can be used to express various words present in English Dictionary. There are 19 different hand shapes to communicate in ASL, which helps to form 26 American manual alphabets. Since, there are fewer hand shapes available, some hand shapes convey different alphabets if the orientation of that hand shape is changed. For Example: 'K' and 'P' letters. Hand gestures are also available to express the numbers from

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:40:15 UTC from IEEE Xplore. Restrictions apply.

'0' to '9'. Though, there are no hand gestures for specific nouns and specific terms [3], various facial and hand gesture signs exist to express various English words. ASL's set of 10 numbers ranging from 0 to 9 and 26 alphabets gesture signs of English Alphabets from A to Z are shown in Figure 1.

American Sign Language or ASL is a normal language, which mostly contains the same characteristics as a spoken language. ASL is used as a mode of conversation by using hand and facial gestures as stated earlier. It is one of the main modes of communication in most North Americans with speaking and hearing disabilities and it is used by normal people too. Sign language recognition will make communication easier if one person isn't well-versed with the language.

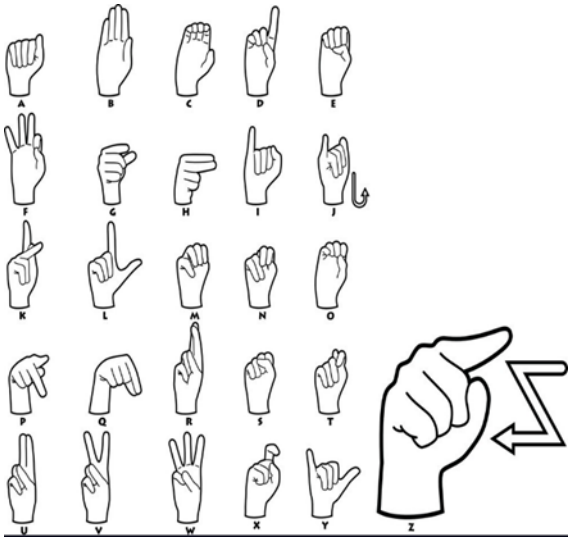


Fig. 1. Poses of Hands for Each Letter of the English Alphabet.

II. LITERATURE SURVEY

Many ways have been put forward to accomplish the problem of recognizing sign language gestures.

Some prior work involves [4] the use of SVM. SVM was earlier used to classify gestures in SASL [5] and, parallel hidden Markov models were used as well.

An approach to classify Sign language is proposed. The model proposed recognizes ASL alphabets gestures with a success rate of 86.67%.

A real-time interaction system between humans and computers based on hand gestures is developed. In this model, images were pre-processed by applying various operations like hand color filtering, morphological transformations, etc. Later the CNN was used on the dataset. A Kalman estimator was used so that whenever a gesture is identified the mouse pointer will move in response to it. Lastly, the system was able to achieve 99.8 % of accuracy on 16 gestures.

Singha made a dataset that contains 240 images of 24 signs present in Indian Sign Language or ISL. Then the Eigenvalues were extracted from the images in the dataset and were classified based on their Euclidean Distance. This system was able to achieve 97% accuracy.

III. METHODOLOGY

Collecting data is one of the most essential and vital parts of any exploration. It is essential to collect data that is relevant and satisfies the requirements needed for the research. The data should contain readings that replicate all scenarios, including extreme cases. This helps in having more thorough and realistic observations. The data collected forms the foundation for any research and should be collected with utmost care.

The proposed research work has decided to design the dataset to match the MNIST classic dataset. Therefore, in the dataset, each element is labelled as training or test cases ranging from 0 to 25, which act as a matched map to all the alphabets in the English language ranging from A-Z. The dataset contains 27,455 training samples and 7172 testing data. This is approximately half the size of the MNSIT dataset, but it replicates it accordingly as every header row has a label of pixel1 to picxel784, which represents the grayscale values of 28x28 pixel images in which the values range from 0-255[13]. The data collected in the MNSIT data is varied by having the same signs made by different subjects in different backgrounds. Moreover, the images were also cropped in various ways; however, the desired hand region was left unchanged.

After exploring the dataset, as shown in figure 2, the data is observed to be evenly balanced across the values. The wide variety of training sample values are evenly distributed, as can be seen from the visualization.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:40:15 UTC from IEEE Xplore. Restrictions apply.

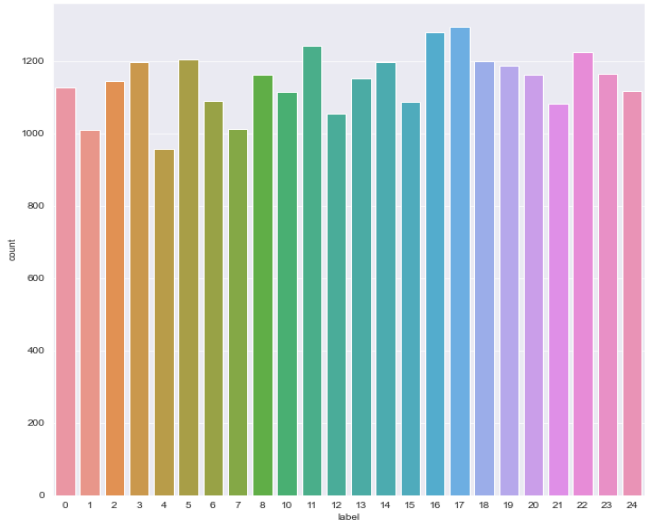


Fig. 2. Number of training examples for each label

A) Loading and Pre-processing of Dataset

We started by importing the dataset to work upon and making it suitable to perform various algorithms to predict sign language alphabet labels correctly. In the given dataset, we already have separate training and test. We visualized our dataset, and it came out to be balanced. We also store the labels of each image, which is mapping with their respective labels in a list, and generates a random image to perform class label verification Figure 3. We performed a grayscale normalization to reduce the effect of illumination's differences. Moreover, the CNN converges faster on [0..1] data than on [0..255] Figure 4. After applying normalization, we split the training set into parts, which is the training set and validation set. So now, we have three sets of data that is training set, validation set, and test set. In this particular problem, each of the 27455 images of the input has 784 columns. To feed this data into the CNN model, we converted the single-dimensional data into CNN accepted format of 3 dimensions. Thus, we accomplished this by converting each row of 784 columns into (28,28,1) matrix.

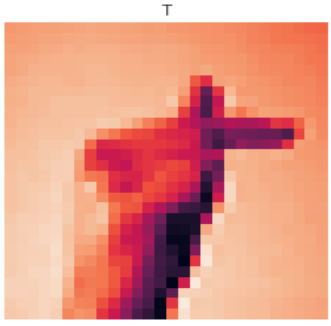


Fig.3 A Random image for class label verification.



Fig. 4. Grayscale images

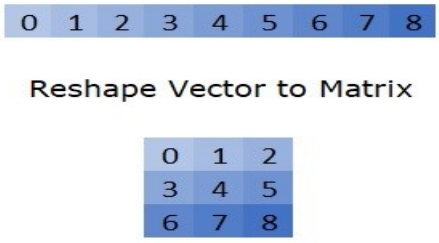


Fig. 5. Reshaping of vector to a matrix.

B) Data Augmentation

In Data augmentation from one image, we generate more images. Therefore, we can have a variety in our dataset, and it is also used to increase the size of the dataset to avoid overfitting. Our data augmentation includes random rotation ranging 10 degrees and random height and width shift in the range of 0.1, random zooming, and random flips. This, we implemented using Keras image pre-processing module.

C) Model Building

The model is developed by using the combination of convolution, max pooling, and flattening layers in Keras. We also used batch normalization and dropout layers to prevent the model from overfitting the data. A callback is used to determine whether the validation set accuracy did not increase even after 2 (patience level) consecutive epochs during the model fitting stage. Then, the learning rate will be altered by a factor of 0.5.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:40:15 UTC from IEEE Xplore. Restrictions apply.

D) CNN Model

The proposed model is sequential since this deep learning model is initialized as a sequence of layers. The information is propagated from the input layer to the hidden layer to the output layer through the model. This network employs a mathematical operation called convolution. The primary purpose of convolution is to find features in our image using a feature detector and put them into a feature map and having them in a feature map, and it preserves the spatial relationship between pixels, which is very important for us.

The first step towards building a model for our work is to initialize the model. Since our model is sequential, we initialized it as sequential. The next step is to create our first layer of CNN to perform convolution operation. We created the input layer of our CNN. We have taken 32 feature detectors of 3x3. The layer takes input into a batch of images. In our case, the batch size is 32. The shape of our image for the input layer is (28,28,1), where the height and width of the image are 28, and there is one channel as the images are greyscaled. Each feature detector produces a feature map after convolution operation over the image with an output volume of 26x26x1. We have taken the default value of padding that is 0 and stride as 1. After the convolution operation, the ReLU activation function (Figure 6.) is applied to increase non-linearity in our CNN model. If the value is greater than 0, it passes the input; otherwise, it returns 0.

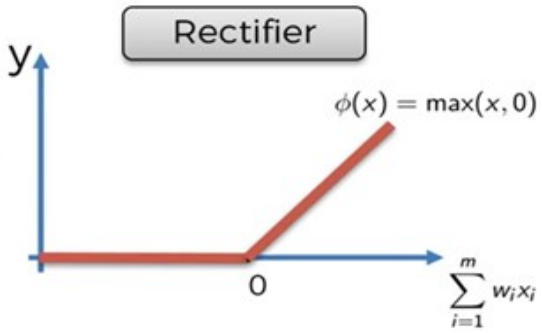


Fig. 6. ReLU Activation Function

Now we have performed Batch Normalization to keep the output of the layer of the same range to get an unbiased result. After this, the output of our first convolution layer of volume 26x26x32 is passed for the max-pooling operation to the next layer. Max pooling help to get rid of unnecessary features and account for their spatial or textual or any kind of distortion. Pooling also helps in preventing any kind of overfitting as it avoids few parameters while performing a pooling operation to get a pooled matrix. Max pooling is used to focus only on the relevant information from the image by extracting only higher values from a portion of input by using an empty feature detector. Pooling is also called downsampling. We are taking stride as two and filter feature detector size as 2x2. As a result, the height and width are reduced by a factor of 2. Thus, the output volume will be 13x13x32. Now we have a second layer, which consists of 1 convolution 2d layer and one max pooling. The input of this layer is the output of the previous layer, which is 13x13x32. The kernel size is 3x3, with 64 different feature detectors. So, the output volume is 11x11x64. Now we again applied Batch Normalization and passed the output volume 11x11x64 as input for Max Pooling, and we get the output as 5x5x64. Now we have the last layer, which is the same as all previous layers consisting of 1 conv2d layer and one max pooling layer [6]. So, the operations performed will be similar, but in the last layer, the number of feature detectors we took is 128. So, the output of the final layer will be 1x1x128[8]. Now we will perform a flattening operation to make it suitable for ANN's input. So, after flattening, it will be converted into a 128x1 vector, and it will be fed to the dense layer (Hidden layer in ANN) with some initialized weights and bias. The output of the dense layer is passed through the ReLU activation function to increase the non-linearity in the model. Now we have used a dropout layer with a dropping rate of 25%, which means 25% of random neurons are switched off so that we can prevent overfitting of our model [12]. Now the output of this hidden or dense layer is fed to the final or output dense layer with 25 neurons. This layer has 25 neurons representing 25 classes. The activation function here we have used is the Softmax activation function. Here we could have used the sigmoid function as the sigmoid function is the heart of the probabilistic approach, but the sigmoid function is good only when we classify between 2 classes. So, when there are more classes, softmax is used, and softmax ensures that the sum of the probability of the output is equal to one, which makes it easy to understand.

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$

Fig. 7. An equation to calculate Softmax Function.

Softmax converts the linear input data into a probability array of all 25 classes, and the probabilities are checked with the actual output, and the loss is calculated using sparse categorical cross-entropy [7]. Here we have used sparse categorical cross-entropy

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:40:15 UTC from IEEE Xplore. Restrictions apply.

because we have labelled as targets.Finally, we compile our model, and to minimize our cost function, The 'Adam' optimizer (stochastic Gradient Descent Method) is used to attain a global minimum. 'Adam' optimizer is used with a learning rate of 0.001 and values of b1=0.9 and b2=0.999. During the backpropagation in the model, weights in the dense layer, as well as weights present in the feature detector changes using the local gradient if not useful for the model. The total trainable parameters to be calculated are 171,993, as seen in Figure 8, representing each layer with its output shape and trainable parameters.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 26, 26, 32)        320
_________________________________________________________________
batch_normalization (BatchNo (None, 26, 26, 32)        128
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 13, 13, 32)        0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 11, 11, 64)        18496
_________________________________________________________________
dropout (Dropout)            (None, 11, 11, 64)        0
_________________________________________________________________
batch_normalization_1 (Batch (None, 11, 11, 64)        256
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 5, 5, 64)          0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 3, 3, 128)         73856
_________________________________________________________________
batch_normalization_2 (Batch (None, 3, 3, 128)         512
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 1, 1, 128)         0
_________________________________________________________________
flatten (Flatten)            (None, 128)               0
_________________________________________________________________
dense (Dense)                (None, 512)               66048
_________________________________________________________________
dropout_1 (Dropout)          (None, 512)               0
_________________________________________________________________
dense_1 (Dense)              (None, 25)                12825
=================================================================
Total params: 172,441
Trainable params: 171,993
Non-trainable params: 448
```

Fig. 8. Trainable Parameters

E) Independent and Dependent Variables

Independent variable- number of images which is 27455, pixel distribution of images which is stored in NumPy array [10].

Dependent variables are- feature map after each convolution layer output, the volume of feature map which is input to the next layer, probabilities of labels which is used for classification.

F) Parameters and Hyper Parameters

Parameters like kernel size, astride of filter, padding, number of features/filters, filter size, activation function, a connection between two layers, pooling type, and number of labels [11].

Hyperparameters like coefficients in the learning rate, dropout rate, batch size, number of hidden & dense layer in architecture, and number of epochs.

G) Performance Measure

Performance measures are used to check the correctness of our model. The majors we are going to use are Accuracy on both training and validation dataset, loss in both validation and training dataset, Precision, Recall, and F1 measure of all 25 classes.

H) Validation Method

We are using the Holdout validation technique to evaluate our model. We initially have a training and testing set. Our dataset contains 34,627 images. 20% of total images (7172) are present in our testing set and the rest in the training set [9]. We split our training set into two parts, one for training our model and another part which is a validation set for validating the proposed model [12-14]. Later, the proposed model is analyzed based on predictions predicted by our trained model on the test set and obtain its accuracy, precision, recall, F1 Score and AUC Score to see its performance.

IV. RESULT AND ANALYSIS

Once the network was trained, it was tested on the testing dataset, including 7,712 samples that were not seen by the network at the time of training, after running it against the testing dataset, performance measures like accuracy, precision, recall, AUC score, and f1 scores were measured, Table1 summarizes the results.

We trained our model on 20 epochs with a batch size of 512. The maximum accuracy obtained on the training set was 99.60%, and invalidating set was 99.95%—epochs. The accuracy graph is shown in Figure 9 and Figure 10 for training and validation accuracy, training and validation loss, respectively.

	Accuracy	Precision	Recall	AUC Score	F1 Score
Result	0.9963	0.99	1.00	0.9981	1.00

Table 1. Accuracy, Precision, Recall, AUC Score, and F1 Score obtained from the test

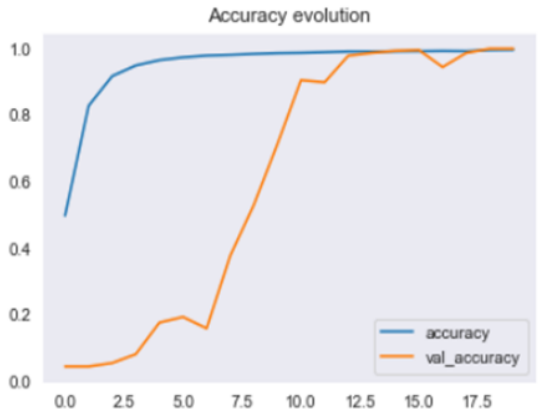


Fig. 9 Accuracy evolution

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:40:15 UTC from IEEE Xplore. Restrictions apply.

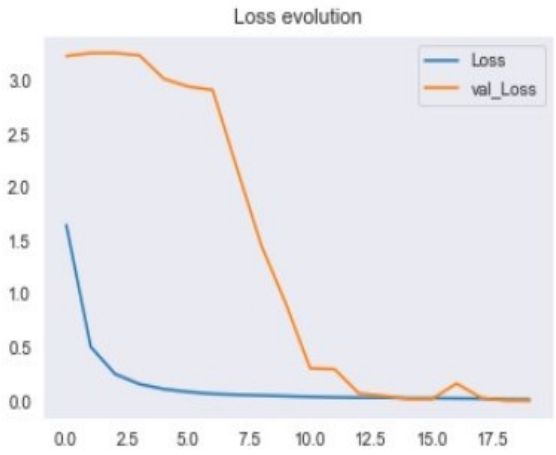


Fig. 10 loss evolution

Along with the calculated performance measures as shown in table 1, a confusion matrix is also obtained. A confusion matrix gives a depiction of the true positives and negatives obtained after running the model. This gives a better understanding as to where the model was faulty and how many true negatives or false positives are there. Figure 11 illustrates this relation with the x-axis showing the predicted levels and the y axis shows the true labels.

The confusion matrix helps in deducing that most of the predictions made by the model are correct. However, some errors can be suspected due to the similarity in some letters. For instance, the fingers' positioning is very similar for O and S, which leads to some faulty predictions. Since we are using the MNIST dataset, most of the predictions are correct. Some of the predictions of our network are shown in Figure 12.

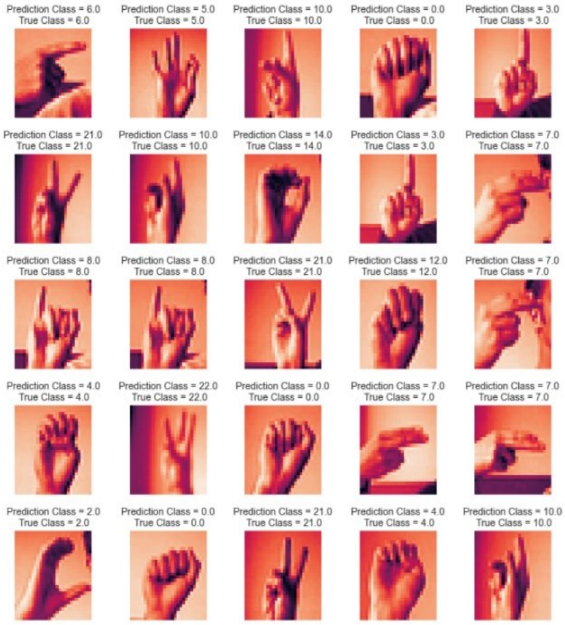


Fig. 12. Predictions of the network on the test set.

The classification report obtained at the end is shown in Figure 13. And precision, Recall, F1 Score for each label is also shown.

	precision	recall	f1-score	support
0.0	1.00	1.00	1.00	331
1.0	1.00	1.00	1.00	432
2.0	1.00	1.00	1.00	310
3.0	1.00	1.00	1.00	245
4.0	1.00	1.00	1.00	498
5.0	0.99	1.00	1.00	247
6.0	1.00	0.99	0.99	348
7.0	0.99	1.00	1.00	436
8.0	1.00	1.00	1.00	288
10.0	1.00	1.00	1.00	331
11.0	1.00	1.00	1.00	209
12.0	1.00	1.00	1.00	394
13.0	1.00	1.00	1.00	291
14.0	1.00	0.99	1.00	246
15.0	1.00	1.00	1.00	347
16.0	1.00	1.00	1.00	164
17.0	0.93	1.00	0.96	144
18.0	1.00	1.00	1.00	246
19.0	1.00	0.97	0.99	248
20.0	1.00	1.00	1.00	266
21.0	1.00	0.97	0.98	346
22.0	1.00	1.00	1.00	206
23.0	0.97	1.00	0.99	267
24.0	1.00	1.00	1.00	332
accuracy			1.00	7172
macro avg	0.99	1.00	1.00	7172
weighted avg	1.00	1.00	1.00	7172

Accuracy: 99.63747908533185

Fig. 13. Classification Report

V. CONCLUSION AND FUTURE WORK

This paper exhibits a method based on convolutional neural networks for recognizing the sign language alphabet. The proposed CNN model has the ability to achieve an accuracy of 99.63% on the proposed test dataset. Initially, when only one convolutional layer is utilized and there was overfitting in the dataset. To overcome this issue of overfitting, two more layers are added so that some parameters are dropped and still preserve the essential features of images. Moreover, the dropout layer, Batch Normalization, and data augmentation are also used to overcome the same and achieve a better performance of our network.

The letters like 'J' and 'Z' were not included in the classification due to the hand movement. So, the video frames are required to classify them. The study may be carried out in the proposed work to accept the video frames and classify the letters like 'J' and 'Z' as well. Moreover, a large dataset of Sign Language is required as the available dataset is quite constrained and most of the researchers are carried out on the same by making it more difficult to compare the work of different researchers.

ACKNOWLEDGMENT

The authors are thankful to the central library of Delhi Technological University Delhi, India for providing all the materials and references required for the lab and research work.

REFERENCES

[1] S. Y. Kim, H. G. Han, J. W. Kim, S. Lee, and T. W. Kim, "A hand gesture recognition sensor using

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:40:15 UTC from IEEE Xplore. Restrictions apply.

reflected impulses," *IEEE Sens. J.*, vol. 17, no. 10, pp. 2975–2976, 2017, doi: 10.1109/JSEN.2017.2679220.

[2] Y. Cui and J. Weng, "A learning-based prediction-and-verification segmentation scheme for hand sign image sequence," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 8, pp. 798–804, 1999, doi: 10.1109/34.784311.

[3] A. Kumar and A. Kumar, "Dog Breed Classifier for Facial Recognition using Convolutional Neural Networks," pp. 508–513, 2020.

[4] K. L. Bouman, G. Abdollahian, M. Boutin, and E. J. Delp, "A low complexity sign detection and text localization method for mobile applications," *IEEE Trans. Multimed.*, vol. 13, no. 5, pp. 922–934, 2011, doi: 10.1109/TMM.2011.2154317.

[5] J. Wu, L. Sun, and R. Jafari, "A Wearable System for Recognizing American Sign Language in Real-Time Using IMU and Surface EMG Sensors," *IEEE J. Biomed. Heal. Informatics*, vol. 20, no. 5, pp. 1281–1290, 2016, doi: 10.1109/JBHI.2016.2598302.

[6] C. Zhang, W. Ding, G. Peng, F. Fu, and W. Wang, "Street View Text Recognition With Deep Learning for Urban Scene Understanding in Intelligent Transportation Systems," *IEEE Trans. Intell. Transp. Syst.*, pp. 1–17, 2020, doi: 10.1109/tits.2020.3017632.

[7] M. Safeel, T. Sukumar, K. S. Shashank, M. D. Arman, R. Shashidhar, and S. B. Puneeth, "Sign Language Recognition Techniques- A Review," *2020 IEEE Int. Conf. Innov. Technol. INOCON 2020*, pp. 1–9, 2020, doi: 10.1109/INOCON50539.2020.9298376.

[8] W. Aly, S. Aly, and S. Almotairi, "User-independent american sign language alphabet recognition based on depth image and PCANet features," *IEEE Access*, vol. 7, pp. 123138–123150, 2019, doi: 10.1109/ACCESS.2019.2938829.

[9] S. Adhikari, S. Thapa, and B. K. Shah, "Oversampling based Classifiers for Categorization of Radar Returns from the Ionosphere," *Proc. Int. Conf. Electron. Sustain. Commun. Syst. ICESC 2020*, no. Icesc, pp. 975–978, 2020, doi: 10.1109/ICESC48915.2020.9155833.

[10] S. Hayani, M. Benaddy, O. El Meslouhi, and M. Kardouchi, "Arab Sign language Recognition with Convolutional Neural Networks," *Proc. 2019 Int. Conf. Comput. Sci. Renew. Energies, ICCSRE 2019*, pp. 2019–2022, 2019, doi: 10.1109/ICCSRE.2019.8807586.

[11] Y. Li, X. Wang, W. Liu, and B. Feng, "Pose Anchor: A Single-Stage Hand Keypoint Detection Network," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 7, pp. 2104–2113, 2020, doi: 10.1109/TCSVT.2019.2912620.

[12] X. Shen and F. L. Chung, "Deep network embedding for graph representation learning in signed networks," *arXiv*, vol. 50, no. 4, pp. 1556–1568, 2019.

[13] Y. Zhu, M. Liao, M. Yang, and W. Liu, "Cascaded Segmentation-Detection Networks for Text-Based Traffic Sign Detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 1, pp. 209–219, 2018, doi: 10.1109/TITS.2017.2768827.

[14] B. K. Shah, V. Kedia, R. Raut, S. Ansari, and A. Shroff, "Evaluation and Comparative Study of Edge Detection Techniques," vol. 22, no. 5, pp. 6–15, 2020, doi: 10.9790/0661-2205030615.

[15] B. K. Shah, V. Kedia and R. K. Jha, "Integrated Vendor-Managed Time Efficient Application to Production of Inventory Systems," 2021 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2021, pp. 275-280, doi: 10.1109/ICICT50816.2021.9358504.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:40:15 UTC from IEEE Xplore. Restrictions apply.

J Mater Sci: Mater Electron

Significant improvements in dielectric, impedance, multiferroic and magnetoelectric properties of $(1 - x)Co_{0.5}Ni_{0.5}Fe_2O_4 - xBaTiO_3$ bulk composites (x = 0, 0.10 and 0.20)

S. Shankar[1,2], O. P. Thakur[3,*], and M. Jayasimhadri[2,*]

[1] *Experimental Research Laboratory, Department of Physics, ARSD College, University of Delhi, Dhaula Kuan, New Delhi 110021, India*

[2] *Luminescent Materials Research Lab, Department of Applied Physics, Delhi Technological University, Bawana Road, Delhi 110042, India*

[3] *Materials Analysis and Research Laboratory, Department of Physics, NSUT, Dwarka, New Delhi 110078, India*

Received: 7 January 2021
Accepted: 18 May 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

ABSTRACT

Magnetoelectric bulk composites of $Co_{0.5}Ni_{0.5}Fe_2O_4$–$BaTiO_3$ (CNFO–BT) were synthesized employing solid-state reaction method. The structural properties of CNFO–BT composites as discussed by X-ray diffraction method confirm lattice distortion and enlarged strain owing to BT substitution in CNFO. The dielectric and impedance measurements exhibit conventional Maxwell–Wagner polarization and confirm the existence of grain dominated non-Debye relaxation phenomena in CNFO–BT composites. The magnetic hysteresis curves reveal strong ferromagnetic behavior in all composites. The maximum energy storage density and an efficiency achieved for 0.8CNFO–0.2BT composite are 4.25 mJ/cm^3 and 31.6%, respectively. The variation of polarization with magnetic field confirmed the highest magnetoelectric coefficient of 5 mV/cm/Oe for 0.8CNFO–0.2BT composite. The variation of dielectric permittivity and ferroelectric polarization with magnetic field reveals lattice distortion, interfacial charge polarization and restricted ferromagnetic domain wall rotation arising from substitution of BT in CNFO. These structure-dependent results suggest potential application of CNFO–BT composites in magnetoelectric sensors and energy storage devices.

1 Introduction

Multiferroics are the materials possessing two or more ferroic orderings, viz. ferroelectric, ferroelastic, ferromagnetic or ferrotoroidic [1]. The magnetoelectric effect in multiferroics materials is ascribed by the magnetic induction of electric polarization and vice versa [2, 3] which makes them suitable for versatile applications such as magnetic memory devices, energy storage and magnetoelectric

Address correspondence to E-mail: opthakur@yahoo.com; jayaphysics@yahoo.com

https://doi.org/10.1007/s10854-021-06227-4
Published online: 02 June 2021

resonator [1, 4]. The extent of this coupling is limited in single-phase materials due to disparate origins of ferroic orderings [5]. Alternatively, the multiferroic magnetoelectric composites are engineered artificially by mixing separately ferroelectric and ferromagnetic phases for enhancing the magnetoelectric coupling at room temperature [5, 6]. Furthermore, the extent of coupling can be maximized by involving piezoelectric and/or piezomagnetic materials which acquaint in achieving strain-mediated coupling and good connectivity at the electric and magnetic phase [5].

Piezomagnetic spinel cobalt ferrite $CoFe_2O_4$ is widely studied material exhibiting strong physiochemical properties owing to its statistical cationic distribution at octahedral/tetrahedral sites and magnetostriction [4, 7]. The unit cell of $CoFe_2O_4$ (CFO) consists of face-centered cubic structure where 8 out of 64 and 16 out of 32 tetrahedral sites and octahedral sites are occupied by cobalt and iron ion, respectively. Due to their high electrical resistance and chemical stability, they are impeccably useful in magnetic storage devices, high-frequency circuits, waveguides and gas sensors [8]. In order to examine the enhanced multiferroic magnetoelectric coupling and to induce the electric nature inside the magnetic spinel ferrite CFO, there is a need of some modifications. Significant efforts for improving the properties of CFO have been focused on partially substituting the Fe^{3+} ions [9, 10], composites with ferroelectric/piezoelectric materials [11], nanoparticles [12] and core/shell structures [13]. Among existing multiferroic magnetoelectric systems, $CoFe_2O_4$−$BaTiO_3$/$PbTiO_3$ composite systems are extensively reconnoitered [2, 4, 6, 14]. The lead-based composites show enhanced electrical properties; however, it is eco-friendly to use lead-free $BaTiO_3$-based composites [15, 16]. $BaTiO_3$ (BT) is the most appropriate suitable substitute for $PbTiO_3$ due to its versatile polymorphic structure and excellent electrical properties [17, 18].

The properties of CFO are, however, always preparation as well as grain size dependent. CFO is hard ferrite, whereas $NiFe_2O_4$ is soft ferrite [19]. The substitution of Ni^{2+} in the cubic structure of CFO is exciting as it alters the electrical properties and magnetic properties due to its lowered magnetic moment in comparison with Co^{2+} [20, 21]. This substitution renders Ni-CFO useful for vast applications with improved magnetoelectric properties [22]. Theoretically strong ME coupling can be achieved if magnetocrystalline anisotropy in CFO is lowered [23]. The recent studies on $Co_{0.5}Ni_{0.5}Fe_2O_4$ (CNFO) have revealed that strong magnetostatic interactions arising from different magnetocrystalline anisotropy are countered by increased coercivity and decreased magnetic saturation [6, 24, 25]. In addition, there have been investigations on $Co_{0.5}Ni_{0.5}Fe_2O_4$-based composites for existence of magnetoelectric properties [1, 4, 14, 26]. The system of CNFO–BT has shown some promising results in magnetic and magnetoelectric properties [22, 27–30]. Nonetheless, the origin of magnetodielectric, impedance and ferroelectric properties in $Co_{0.5}Ni_{0.5}Fe_2O_4$–$BaTiO_3$ composites with $Co_{0.5}Ni_{0.5}Fe_2O_4$ as host material is not explored thoroughly. These composites are expected to bring improvements in electrical and magnetoelectric properties in CNFO.

In this manuscript, composites of ferromagnetic $Co_{0.5}Ni_{0.5}Fe_2O_4$ and lead-free ferroelectric $BaTiO_3$ are prepared via solid-state reaction method. The effect of $BaTiO_3$ on dielectric, impedance, multiferroic and magnetoelectric properties of $(1-x)Co_{0.5}Ni_{0.5}Fe_2O_4$–$xBaTiO_3$ bulk composites (x = 0, 0.10 and 0.20) has been studied extensively. These results are beneficial for exploiting lead-free CNFO–BT composites for magnetoelectric and energy storage applications.

2 Experimental

2.1 Synthesis of $Co_{0.5}Ni_{0.5}Fe_2O_4$ and $Co_{0.5}Ni_{0.5}Fe_2O_4$–$BaTiO_3$ composites

Bulk $Co_{0.5}Ni_{0.5}Fe_2O_4$ and $(1-x)Co_{0.5}Ni_{0.5}Fe_2O_4$–$xBaTiO_3$ composites were prepared using high purity powders of Co_3O_4, Fe_2O_3, NiO, $BaCO_3$ and TiO_2 employing solid-state reaction method. Stoichiometric amounts of chemicals for $Co_{0.5}Ni_{0.5}Fe_2O_4$ (CNFO) and $BaTiO_3$ (BT) were weighed mixed and grounding agate pestle mortar for 6 h. The mixture was calcined at 1000 °C for 12 h in air. The powders of $Co_{0.5}Ni_{0.5}Fe_2O_4$ and $BaTiO_3$ were stoichiometrically mixed for preparing compositions $(1-x)Co_{0.5}Ni_{0.5}Fe_2O_4$–$xBaTiO_3$ with (x = 0, 0.10, 0.20) and named as CNFO ($Co_{0.5}Ni_{0.5}Fe_2O_4$), CNFO–10BT ($0.9Co_{0.5}Ni_{0.5}Fe_2O_4$–$0.1BaTiO_3$) and CNFO–20BT ($0.8Co_{0.5}Ni_{0.5}Fe_2O_4$–$0.2BaTiO_3$). The resultant powders were pressed into 10-mm-diameter pellets using polyvinyl alcohol as binder and sintered at 1200 °C for 8 h along with

calcined powder of respective sample to avoid oxygen loss in sintering process.

2.2 Characterization techniques

The structural analysis has been carried out using Rikagu X-ray diffractometer employing Cu-K_α radiation with wavelength of 1.5418 Å. The XRD data have been recorded from the powder obtained by crushing the sintered pellets. The pellets of CNFO–BT composites were silver coated to improve electrical contact. Dielectric and impedance measurements of the coated composite pellets were made using Wayne Kerr 6500 P high-frequency LCR meter setup in the frequency and temperature range of 100-1 MHz and 100–400 °C, respectively. The ferroelectric *P–E* hysteresis loop for the coated pellets was studied using Marine India *P–E* loop tracer setup operating at 50 Hz with electric field up to 5 kV/cm and magnetic field up to 4000 Oersted.

3 Results and discussion

The X-ray diffraction patterns of CNFO and CNFO–BT composites are shown in Fig. 1. The XRD plot of CFO is referred from our earlier work [4]. The XRD pattern of CNFO is matched with JCPDS # 22-1086 and confirms the spinel cubic structure with space group (*Fd3m*). The splitting and broadening of peaks in XRD of CNFO in comparison with XRD peaks of CFO at $2\theta \sim 30°$ and 62.5° are matched with JCPDS # 10-0325 for $NiFe_2O_4$ [31] and are marked by '*' symbol in Fig. 1. This occurrence of $NiFe_2O_4$ peaks confirms the formation of CNFO. The impurities in XRD of CNFO between 2θ of 50° and 60° are due to presence of Fe_2O_3 and are matched with JCPDS # 02-0919. The XRD patterns of CNFO–BT system are matched with JCPDS # 22-1086 for CNFO and JCPDS # 05-0626 for $BaTiO_3$ (BT) and confirm the formation of CNFO–BT composites. The ferroelectric BT phase in the composite at $2\theta \sim 32°$ and at other higher values is indicated by '#' symbol in Fig. 1.

The effect of BT addition in CNFO is observed by studying the lattice parameter variation and lattice strain estimations. The variation of lattice parameter was obtained using Debye–Scherrer equation [32] and lattice strain was obtained from Williamson–Hall equation [33], and the subsequent results are plotted in Fig. 2. It can be observed that the lattice parameter

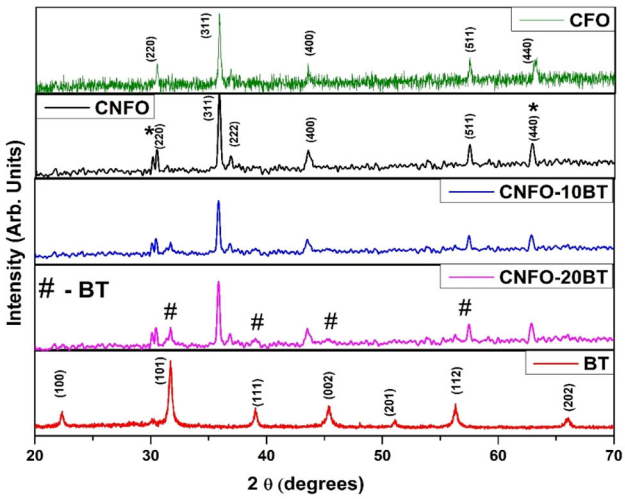


Fig. 1 XRD patterns of CNFO and CNFO–BT composites

obtained for CNFO is 8.325 Å which is lower than that of $CoFe_2O_4$ (10.7 Å) [4] as at both octahedral/tetrahedral sites, ionic radii of Ni^{2+} (0.745/0.55 Å) are less than Co^{2+} (0.69/0.58 Å) [34]. The lattice parameters in Fig. 2 decrease from 8.325 to 7.967 Å for CNFO to CNFO–BT composites due to lattice distortion arising on account of trapping centers formed by BT substitution. This lattice distortion is also accompanied by the increase in the lattice strain as shown in Fig. 2.

The dielectric analysis is an ideal technique for correlation of most of the properties of ferrites with preparation, composition and temperature. The frequency dependency of dielectric constant for CNFO and CNFO–BT composites is shown in Fig. 3a–c. The graphs show usual behavior of dielectric materials

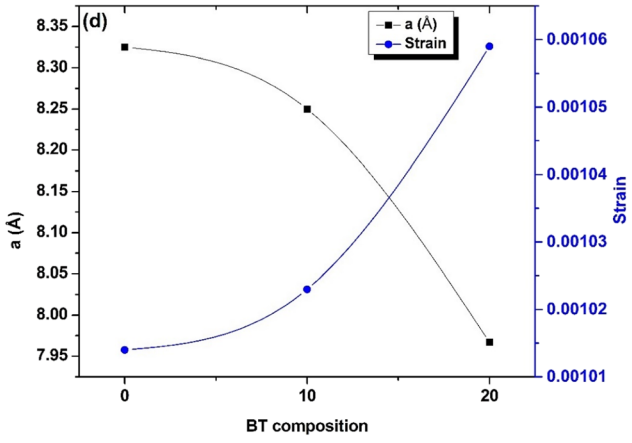


Fig. 2 Variation of lattice parameter and strain in CNFO–BT composites with BT composition

with CNFO and CNFO–BT composites possessing large dielectric constant values [35] at low frequencies, and the dielectric constant decreases at higher frequencies.

The polarization mechanism of any dielectric material can be of the form of atomic, electronic, dipole and interface [36, 37]. The first two mechanisms, viz. electronic and atomic polarizations, usually have instantaneous effects and are observed at high frequencies only [36, 37]. The dipolar polarization in materials possessing polar groups is observed as step-like decrease in dielectric permittivity with frequency. The main reason to show such a decreasing trend is due to the presence of interfacial polarization as predicted by Maxwell–Wagner [38] and arises because of the charge accumulation on sample–electrode interface as well as on the interface between the constituent phases [36, 37]. The exchange of electrons among multivalent Fe^{2+}/Fe^{3+} creates a spatial displacement of Fe ions along with electric field resulting in the polarization. The high value of dielectric permittivity at low frequencies is because of the presence of multivalent Fe ions, its related oxygen vacancies and defects. The decrement in the value of dielectric permittivity with frequency is because the dipoles cannot follow the electric field oscillations at higher fields. As per Koops theory, the contribution of dielectric constant at lower frequency originates from grain boundaries as they possess high impedance at grain boundary region [39]. The contribution to dielectric permittivity at high frequencies originates from low resistive grains. Thus, the presence of

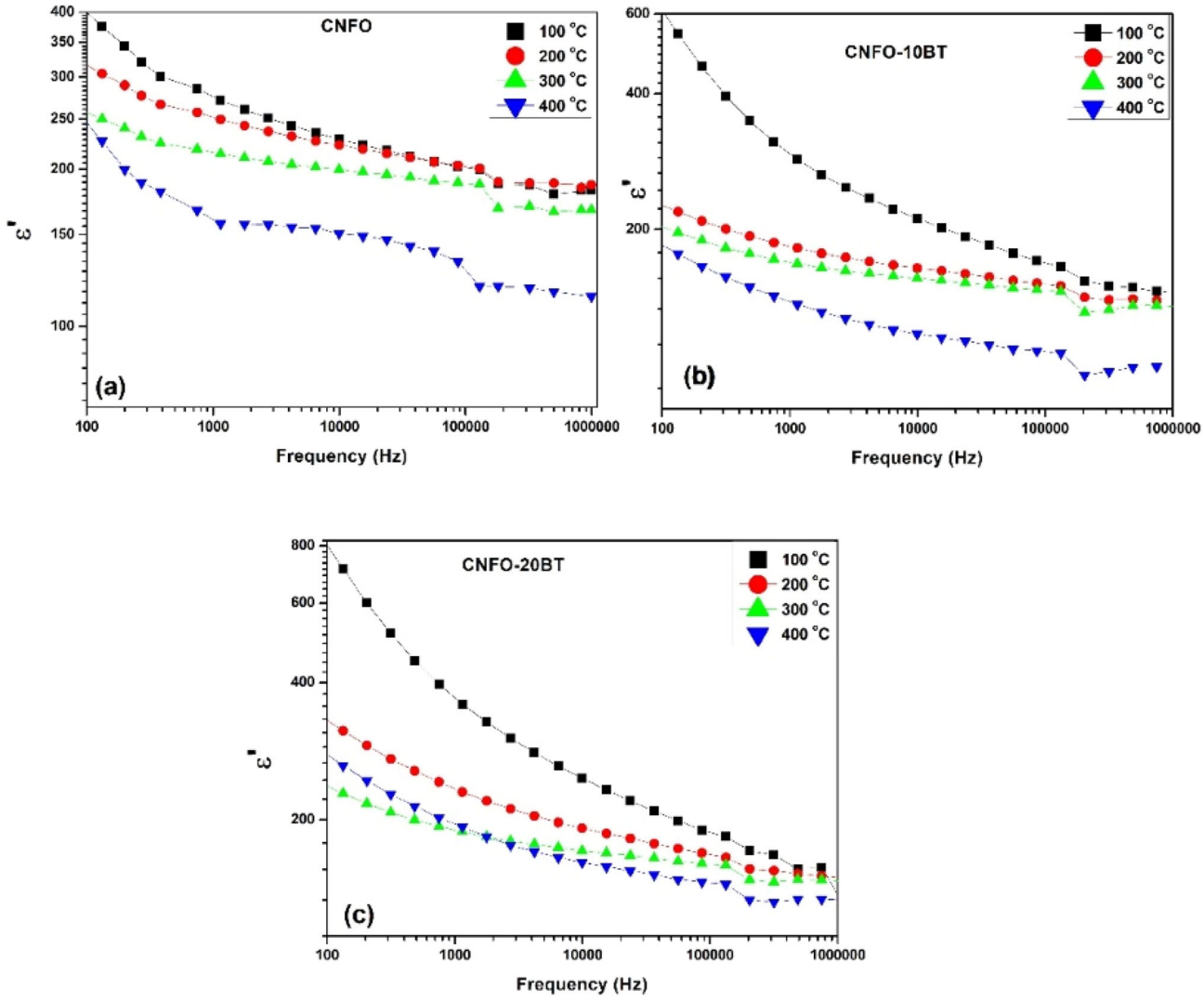


Fig. 3 Frequency dependence of real part of dielectric constant (ε') for CNFO and CNFO–BT composites at different temperatures

multivalent Fe ions is responsible for polarization in CNFO. The dielectric constant increases with BT content and suggests enhancement in electrical nature of CNFO–BT composites. For all composites, the increase in temperature from 100 to 400 °C onsets the space charge polarization linked to ionic defect movements and decreases the dielectric constant [40].

The magnetodielectric coupling constant (MD) is studied by measuring the dielectric properties at different magnetic fields and at a frequency of 10 kHz as shown in Fig. 4 and is quantified as

$$\mathrm{MD} = \frac{\epsilon'(H) - \epsilon'(0)}{\epsilon'(0)} \times 100 \quad (1)$$

The composition of magnetic–ferroelectric phase is an important factor for deciding the magnetodielectric response. The variation in dielectric constant is less for CNFO and significantly increases with BT content. This enhancement could be possibly due to the modification of spatial charge distribution and restricted ferromagnetic domain wall rotation originating from BT substitution in CNFO. With increasing magnetic field, the change in dielectric constant drastically decreases due to alignment of domain walls in CNFO. Surprisingly, all CNFO–BT composites indicate stability in magnetodielectric coupling by displaying non-zero MD value at even higher fields. It is important to note that the magnetodielectric response in CNFO as well as CNFO–BT composites originates due to the contribution of magnetostriction of CNFO. This occurs due to the strain induced in CNFO phase on application of magnetic field creating a mechanical coupling between dielectric and magnetic phases and can be expressed as strain-mediated magnetodielectric response. Additionally, the interfacial polarization effect in CNFO–BT can also induce magnetodielectric behavior based on the impedance of grain or grain boundary with application of magnetic field [41].

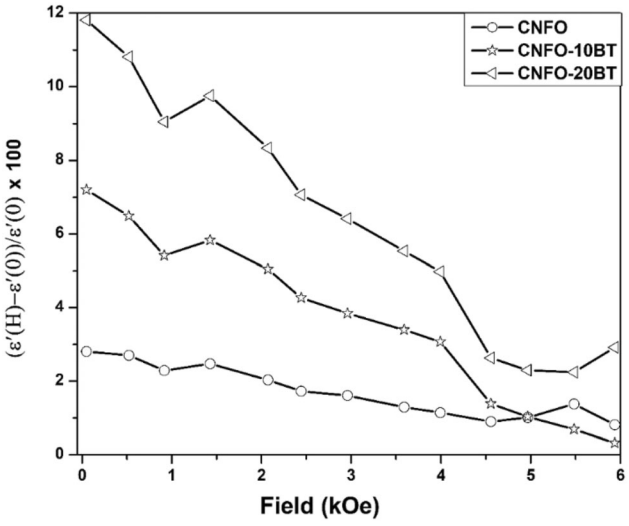


Fig. 4 Percentage change in dielectric constant for CNFO and CNFO–BT composites as a function of magnetic field at 10 kHz

Figure 5 shows Nyquist plots of CNFO and CNFO–BT composites. All the composites display single semicircular arc starting from the origin whose radius decreases with temperature suggesting dominance of grains contribution and negative temperature coefficient of resistance [42]. It is noted that the impedance value in the composites increases with BT suggesting highly resistive nature of CNFO–BT composites [43]. CNFO is a weak ferroelectric and poor dielectric material. Thus, the significant contribution in the dielectric and impedance properties is due to the effective content of BTO. Increasing the amount of BTO in CNFO–BT composite therefore increases the net impedance of CNFO–BT composite [43]. The centers of semicircles are observed to be positioned below the Z'-axis owing from spread relaxation times and indicate the existence of non-Debye relaxations in CNFO–BT composites [44]. The non-Debye relaxations can be understood by representing the Nyquist plots in terms of modified R–C electric circuits. These electric circuits contain arrangement of one or more subarray of resistance (R), capacitance (C) and constant phase element (CPE) with each subarray denoting the contributions of grains and grain boundary. The non-Debye behavior present in BT-CFO composites is justified from the presence of CPE in the equivalent circuit, which is related to the impedance by the equation $Z_{\mathrm{CPE}} = 1/(j\omega)^n$ CPE. The corresponding fitted parameters are displayed in Table 1. The impedance spectra of CNFO–BT composites are fitted with one R-CPE parallel combination and are in agreement with each other. The fitting also confirms the dominance of grain resistance in the conduction phenomena.

Figure 6 shows magnetic hysteresis curves for CNFO and CNFO–BT composites. Inset of Fig. 6 shows the enlarged view of M-H curves at low magnetic fields. All composites display systematically strong ferromagnetic behavior. The saturation magnetization for CNFO is comparable to the earlier reported work [25]. However, the saturation

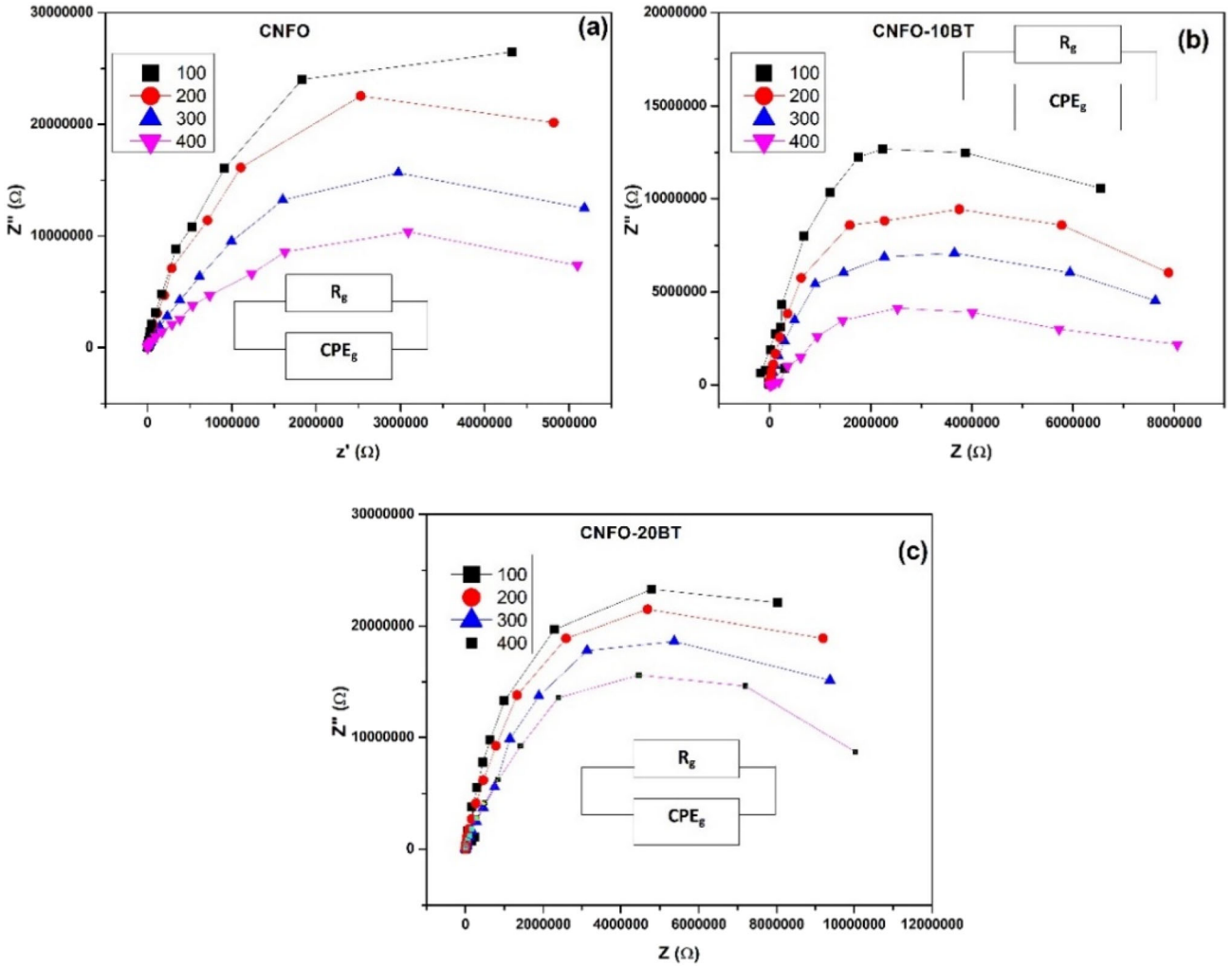


Fig. 5 Nyquist plot of CNFO–BT composites measured at different temperatures along with their equivalent circuit measured

Table 1 Fitted parameters for Cole–Cole plots of CNFO–BT composites

Temp (°C)	R_g (Ω)	CPE_g	n_g
CNFO			
200	6.1×10^6	2×10^{-8}	0.65
300	4.3×10^6	6×10^{-9}	0.78
400	1.9×10^6	9×10^{-10}	0.86
CNFO–10BT			
200	8.2×10^6	1.2×10^{-9}	0.62
300	5.3×10^6	5×10^{-9}	0.70
400	3.1×10^6	7×10^{-10}	0.79
CNFO–20BT			
200	8.9×10^6	3×10^{-8}	0.55
300	5.4×10^6	7×10^{-8}	0.69
400	2.9×10^6	4×10^{-9}	0.75

magnetization is smaller than CNFO prepared by different methods [45, 46] which could be due to less duration mixing of primary constituents resulting in incomplete homogeneity of the structure [25]. The magnetic retentivity is found to decrease with increase in content of BT. This re-emphasizes structural distortion induced in CNFO lattice owing to BT substitution. Contrarily, the coercivity is large for CNFO–10BT and least for CNFO–20BT as compared to CNFO. The increased value of coercivity in CNFO–10BT composite is due to enhancement in magnetoelectric coupling; however, such coupling is limited by lattice distortion and weakens the ferromagnetic nature in CNFO–BT composite. On further increasing the BT content as in case of CNFO–20BT, enlarged strain-mediated magnetoelectric coupling at the piezoelectric–piezomagnetic interface enhances the ferromagnetic nature.

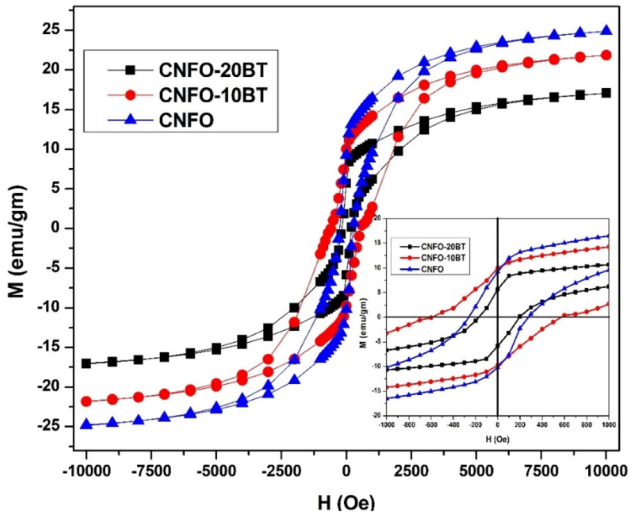


Fig. 6 Room temperature magnetization curves (M vs. H) for CNFO and CNFO–BT composites. Inset shows the enlarged view of CNFO and CNFO–BT composites

The ferroelectric hysteresis P–E loops of CNFO–BT composites are plotted in Fig. 7. All the loops show weak P–E ferroelectric polarization loop because of spatially distributed electric dipoles. The unsaturated lossy P–E hysteresis loop is originated due to the oxygen vacancies and further to high leakage current in the composites. These oxygen vacancies have been created during synthesis of single oxidized ions of Fe and Ti atoms and are expressed by Kroger–Vink notation

$$O_o^x \rightarrow \frac{1}{2}O_2 + V_o^{**} + e^- \quad (2)$$

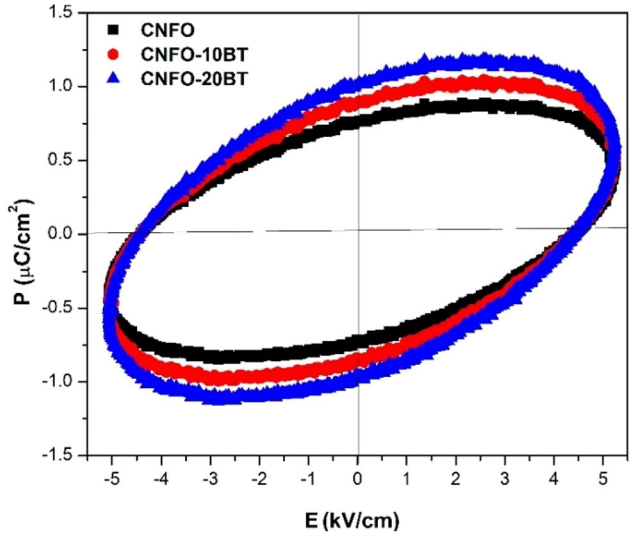


Fig. 7 P–E hysteresis loops of CNFO and CNFO–BT composites

where V_o^{**} is the ionized metal [47]. The magnetic nature of CNFO particles tends to agglomerate and thereby induces weak ferroelectricity [23, 30]. As the content of BT phase is increased in CNFO, the maximum polarization is increased due to increased conductivity. The energy storage and efficiency are estimated from ferroelectric measurements [48]. The unloaded sample of CNFO and CNFO–20BT displayed the energy density of 2.1 and 4.25 mJ/cm^3, respectively, with an efficiency of 17 and 31.6%, respectively.

The magnetoelectricity in CNFO–BT composites can be understood well by examining the variation of maximum electric polarization versus applied electric field hysteresis graphs as a function of external magnetic field [49]. Figure 8 shows the variation in maximum polarization measured from P–E hysteresis loop at different magnetic fields for CNFO and CNFO–BT composites at room temperature with fixed frequency of 50 Hz. The maximum polarization observed decreases with increasing magnetic field suggesting strong magnetoelectric character in CNFO composites. The largest change in maximum polarization of 20% is observed in CNFO–20BT composite, making it most suitable for magnetoelectric devices [50]. This enhanced magnetoelectric effect is correlated with lattice shrinkage on BT substitution resulting in strain-mediated magnetoelectric coupling at the piezoelectric–piezomagnetic interface [26]. In order to obtain the ME coupling coefficient, the variation of P_{max} was linearly fitted with magnetic field H. The error bars obtained from linear fit

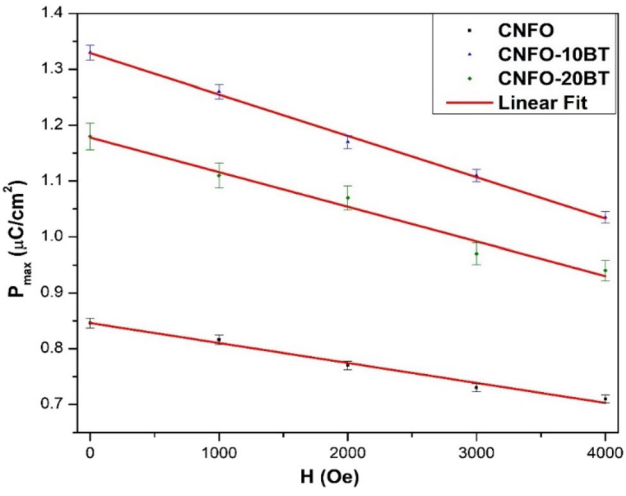


Fig. 8 Variation of maximum polarization (P_{max}) with magnetic field for CNFO and CNFO–BT composites

are also incorporated in the experimental data plot. The maximum magnetoelectric coupling coefficient (α) was achieved for CNFO–20BT composite as 5 mV/cm/Oe. This magnetoelectric response of CNFO–BT composite agrees with some reported results [29, 43].

4 Conclusions

$Co_{0.5}Ni_{0.5}Fe_2O_4$ and $(1-x)Co_{0.5}Ni_{0.5}Fe_2O_4$–$xBaTiO_3$ (CNFO–BT) magnetoelectric composites were synthesized from solid-state reaction route. The XRD studies reveal lattice distortion and increased strain with BT substitution due to formation of trapping centers in CNFO. The dielectric and impedance measurements reveal presence of interfacial polarization and non-Debye relaxations in CNFO–BT composites. The occurrence of ferroelectric and ferromagnetic hysteresis loops establishes the multiferroic nature in CNFO–BT composite. The evidence of magnetoelectric coupling in CNFO–BT composites is confirmed from magnetodielectric coupling and magnetic-dependent ferroelectric measurements. These results are useful for technological advancements of CNFO–BT composites in magnetoelectric device applications.

Acknowledgements

This work is funded by Research Scheme no. 03(1427)/18/EMR-II, CSIR, New Delhi, and supported by USIC, University of Delhi, New Delhi, India, for characterization facilities. The authors are also thankful to Netaji Subhas University of Technology (NSUT), New Delhi, India, for *P*–*E* characterization facility and University Science Instrumentation Centre (USIC), University of Delhi, Delhi, India, for XRD and magnetic characterization facility.

References

1. S. Mohan, P.A. Joy, Ceram. Int. **45**, 12307 (2019)
2. M.D. Chermahini, M.M. Shahraki, M. Kazazi, Mater. Lett. **233**, 188 (2018)
3. Y. Shu, Q. Ma, Z. Ding, L. Cao, X. Chen, F. Yang, Phys. Lett. A **383**, 911 (2019)
4. S. Shankar, M. Kumar, A.K. Ghosh, O.P. Thakur, M. Jayasimhadri, J. Alloys Compd. **779**, 918 (2019)
5. S. Liu, S. Yan, H. Luo, S. Huang, C. Liao, L. Deng, Mater. Lett. **212**, 139 (2018)
6. S. Lather, A. Gupta, J. Dalal, V. Verma, R. Tripathi, A. Ohlan, Ceram. Int. **43**, 3246 (2017)
7. J. Wang, F. Zhao, W. Wu, S. Cao, G. Zhao, Phys. Lett. A **376**, 547 (2012)
8. V.K. Chakradhary, A. Ansari, M.J. Akhtar, J. Magn. Magn. Mater. **469**, 674 (2019)
9. C. Schmitz-Antoniak, D. Schmitz, P. Borisov, F.M.F. de Groot, S. Stienen, A. Warland, B. Krumme, R. Feyerherm, E. Dudzik, W. Kleemann, H. Wende, Nat. Commun. **4**, 2051 (2013)
10. A. Manjeera, M. Vittal, V.R. Reddy, G. Prasad, G.S. Kumar, Ferroelectrics **519**, 15 (2017)
11. M.T. Rahman, C.V. Ramana, J. Appl. Phys. **116**, 164108 (2014)
12. M. Kumar, S. Shankar, O.P. Thakur, A.K. Ghosh, Mater. Lett. **143**, 241 (2015)
13. S. Chandarak, A. Ngamjarurojana, S. Srilomsak, P. Laoratanakul, S. Rujirawat, R. Yimnirun, Ferroelectrics **410**, 75 (2010)
14. M.U.D. Rather, R. Samad, B. Want, J. Mater. Sci. Mater. Electron. **29**, 19164 (2018)
15. N. Ortega, A. Kumar, P. Bhattacharya, S.B. Majumder, R.S. Katiyar, Phys. Rev. B - Condens. Matter Mater. Phys. **77**, 1 (2008)
16. R.K. Singh, S. Sanodia, N. Jain, R. Kumar, AIP Conf. Proc. **1953**, 1 (2018)
17. S. Shankar, O.P. Thakur, M. Jayasimhadri, Mater. Chem. Phys. **234**, 110 (2019)
18. M.P.K. Sahoo, Y. Zhang, J. Wang, Phys. Lett. A **380**, 299 (2015)
19. R. Pandey, L.K. Pradhan, S. Kumar, M. Kar, J. Alloys Compd. **762**, 668 (2018)
20. P. Liu, Z. Yao, V.M.H. Ng, J. Zhou, L.B. Kong, Mater. Lett. **248**, 214 (2019)
21. P. Liu, Z. Yao, J. Zhou, Z. Yang, L.B. Kong, J. Mater. Chem. C **4**, 9738 (2016)
22. Manjusha, K.L. Yadav, J. Mater. Sci. Mater. Electron. **27**, 6347 (2016)
23. G.D. Dwivedi, K.F. Tseng, C.L. Chan, P. Shahi, J. Lourembam, B. Chatterjee, A.K. Ghosh, H.D. Yang, S. Chatterjee, Phys. Rev. B - Condens. Matter Mater. Phys. **82**, 3 (2010)
24. P. Puliová, J. Kováč, A. Voigt, P. Raschman, J. Magn. Magn. Mater. **341**, 93 (2013)
25. M.D. Chermahini, H.A. Baghbaderani, M.M. Shahraki, M. Kazazi, Ceram. Int. **45**, 5491 (2019)

26. D. Zhang, J. Cheng, J. Chai, J. Deng, R. Ren, Y. Su, H. Wang, C. Ma, C.S. Lee, W. Zhang, G.P. Zheng, M. Cao, J. Alloys Compd. **740**, 1067 (2018)
27. N. Pulphol, R. Muanglua, S. Niemcharoen, W. Pecharapa, W.C. Vittayakorn, N. Vittayakorn, Adv. Mater. Res. **802**, 22 (2013)
28. N. Pulphol, R. Muanghlua, S. Niemcharoen, N. Vittayakorn, W. Vittayakorn, Ferroelectrics **488**, 170 (2015)
29. S. Pachari, S.K. Pratihar, B.B. Nayak, RSC Adv. **5**, 105609 (2015)
30. R. Gao, Q. Zhang, Z. Xu, Z. Wang, C. Fu, G. Chen, X. Deng, X. Luo, Y. Qiu, W. Cai, J. Mater. Sci. Mater. Electron. **30**, 10256 (2019)
31. M. Parishani, M. Nadafan, Z. Dehghani, R. Malekfar, G.H.H. Khorrami, Results Phys. **7**, 3619 (2017)
32. B.D. Cullity, *Elements of X-Ray Diffraction*, 2nd edn. (Addison-Wesley Publishing Company Inc, New York, 1978)
33. V. Mote, Y. Purushotham, B. Dole, J. Theor. Appl. Phys. **6**, 6 (2012)
34. N. Kasapoglu, B. Birs, Cent. Eur. J. Chem. **5**, 570 (2007)
35. A. Kumar, P. Sharma, D. Varshney, Ceram. Int. **40**, 12855 (2014)
36. S. Xavier, S. Thankachan, B.P. Jacob, E.M. Mohammed, IOP Conf. Ser. Mater. Sci. Eng. **73**, 012093 (2015)
37. S. Thankachan, B.P. Jacob, S. Xavier, E.M. Mohammed, J. Magn. Magn. Mater. **348**, 140 (2013)
38. R. Debnath, S.K. Mandal, A. Nath, Mater. Lett. **237**, 80 (2019)
39. C.G. Koops, Phys. Rev. **83**, 121 (1951)
40. A. Jain, A.K. Panwar, A.K. Jha, Mater. Res. Bull. **100**, 367 (2018)
41. G. Catalan, Appl. Phys. Lett. **88**, 1 (2006)
42. S. Shankar, M. Kumar, S. Kumar, O.P. Thakur, A.K. Ghosh, J. Alloys Compd. **694**, 715 (2017)
43. R. Gao, Q. Zhang, Z. Xu, Z. Wang, G. Chen, X. Deng, C. Fu, W. Cai, Compos. Part B Eng. **166**, 204 (2019)
44. M. Kumar, S. Shankar, S. Brijmohan, O.P. Kumar, Thakur, A.K. Ghosh, Phys. Lett. A **381**, 379 (2017)
45. Z.H. Low, I. Ismail, M.S.E. Shafie, I.R. Ibrahim, M. Ertuğrul, R.S. Azis, N. Mohd Saiden, I.H. Hasan, F. Mohd, Idris, R. Nazlan, Results Phys. **15**, 102683 (2019)
46. G. Muscas, G. Concas, S. Laureti, A.M. Testa, R. Mathieu, J.A. De Toro, C. Cannas, A. Musinu, M.A. Novak, C. Sangregorio, S.S. Lee, D. Peddis, Phys. Chem. Chem. Phys. **20**, 28634 (2018)
47. S. Shankar, O.P. Thakur, M. Jayasimhadri, J. Electron. Mater. **49**, 472 (2020)
48. S. Shankar, M. Kumar, V. Tuli, O.P. Thakur, M. Jayasimhadri, J. Mater. Sci. Mater. Electron. **29**, 18352 (2018)
49. M. Kumar, S. Shankar, R.K. Kotnala, O. Parkash, J. Alloys Compd. **577**, 222 (2013)
50. M. Sufyan, S. Atiq, S.K. Abbas, M. Younis, Mater. Lett. **238**, 10 (2018)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021)
IEEE Xplore Part Number: CFP21K74-ART; ISBN: 978-0-7381-1327-2

Solar Energy Prediction using Decision Tree Regressor

Rahul
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
rahul@dtu.ac.in

Aakash Gupta
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
aakashg1999@gmail.com

Ankur Bansal
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
ankur.b1999@gmail.com

Kshitij Roy
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
roy.kshitij@gmail.com

2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) | 978-1-6654-1272-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICCS51141.2021.9432322

***Abstract*—The energy demand is increasing every day. This increase in demand has led us to find alternate sources of energy. The non-renewable sources have limited supply, and constant use gives rise to various types of pollution. Gradually, people started relying on renewable sources of energy due to their abundance in nature. Alternately, Solar Energy (SE) is one of the best sources of energy in India. This paper aims to predict the output production of Solar Power Plants in KWh. Also, this research paper attempts to find the relation of weather attributes with the output power generated. The Decision Tree Regressor model will be used to forecast the power generation of a Solar Power Plant. This model's predictions will help us determine how a solar power plant can be efficiently used to generate a high amount of power. The motivation behind the study was to develop a system that can effectively predict the output of these photovoltaic cells so that orders for supply and demand gaps are placed timely, ensuring the quality of service and cost savings.**

Keywords*— *Machine Learning, Solar Energy, Electrical Energy, Decision Tree, Renewable Energy, Regression, Photovoltaic cells.

I. INTRODUCTION

India has witnessed a rapid demand for non-renewable sources of energy in the past several years. Due to their limited supply and increase in pollution levels, it has been gradually attempted to find alternatives to such resources. One of the best alternatives in India is SE, considering its abundance and tropical climatic conditions. SE is generated with the help of photovoltaic cells. These energy sources highly depend on weather conditions, and they always may not be favorable for the production of electrical energy. In such cases, expensive sources of energy will be required to fulfill its shortfall. Thus, a system will be required to effectively predict the output of these photovoltaic cells so that orders for supply and demand gaps are placed timely, ensuring the quality of service and cost savings.

The efficiency of a photovoltaic cell is also a major concern. A maximum of 75% of the SE to earth can be converted to electricity by a silicon semiconductor, but that is not the case since the PV cells are able to give efficiency up to 33% only. Minimum energy is required by the photon to excite an electron, i.e., to remove it from its crystal structure, known as the band-gap energy. The storage of the generated energy is also a major issue as the current battery technology has a long way to go.

Gujarat Power Corporation Limited (GPCL) is Gujarat state's energy generation and distribution company. When faced with the challenge to cater to the ever-increasing energy demands, GPCL launched the "Gujarat Solar Park" project in December 2010. It was inaugurated in April 2012 in Patan district with Rs.4500Cr budget (Rs.3996Cr for power plant and rest for land acquisition and infrastructure). It is currently Asia's largest and Asia's first Multi-Facility Solar Park. This Solar Park accounts for 10% of India's total solar capacity. The "Gujarat Solar Park" has 224 MW of installed Solar Power capacity while 50MW more is expected to be completed soon and 100 MW of Wind Power. Developing a large cluster of SE unit in one place has resulted in savings of 40% of the cost and 3,42,400 tons of Carbon Emission Reduction[1].

Various machine learning and statistical models have been proposed to predict SE generation, each with its own shortcoming, thus leaving scope for further research since forecasting weather in itself is a challenging task. One way to revamp the accuracy and precision of the prediction is by developing efficient algorithms. The other way is to incorporate more and more relevant weather-related and PV-related parameters in our datasets. In this research, we have touched upon the scope of improvement. Our attempt includes

[1] https://gpcl.gujarat.gov.in/showpage.aspx?contentid=15

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:31:34 UTC from IEEE Xplore. Restrictions apply.

the use of a **Decision Tree Regressor** for this regression problem, and our results demonstrate that the level of accuracy obtained is satisfactory. Secondly, to accomplish our objective, we make use of several relevant weather parameters provided in a public forum by Nasa Power Project [2] like the maximum temperature at ground level, precipitation, dew/frost data, wind speed range, surface pressure, etc., in addition to the parameters in our original dataset.

II. PREVIOUS WORKS

Nowadays almost everything depends on the steady supply of electricity. The petroleum derivatives are a vital wellspring of dependable and reliable energy. However, they have significant expenses related to them. They discharge hazardous gases and are dependent upon the vulnerability of the global oil costs. Then again, SE is modest and a clean wellspring of energy that can be created by many countries. However, its reliance on climate conditions makes it less solid. To adapt to this lack of quality, the expectation of SE contributes towards a steady stockpile. The forecast of SE is multidisciplinary research that needs commitment from meteorology, sun-powered cell designing, electrical designing, and AI prediction.

The comparative study of the solar power generation predicting models is mentioned in [1],[2]. Weather data was taken as input to make the predictions in most of the studies conducted. Machine learning methods like Support Vector Regressor (SVR) have been used in conjunction with ensemble technique to further improve on the SVR [3]. Deep learning methods are also being used where we train Neural Networks on our training dataset to yield high accuracy results. Artificial Neural Networks were used with correlation-based feature selection, and compared with other regressive models, the model achieved 2.1436 RMSE (Root Mean Square Error) [4].

The Recurrent Neural Networks have always suffered from the vanishing and exploding gradient problem, which has been solved by modifying them and adding Gate controls; these networks are called LSTM (Long Short-Term Memory). These models were also used to predict SE Prediction [5]. Another type of Neural network, namely, Convolutional Neural Networks (CNN), where we have layers of convolutions and used mostly in computer vision, was used in this research paper [6]. Classification models like K-Nearest Neighbours (KNN) and Support Vector Machine (SVM) are used in Day Ahead Short Term (DAST) PV power prediction for weather classification [7].

Some studies compare these models, such as regression(linear, polynomial), Boosting(XGBoost, AdaBoost, Gradient Boosting), Bagging [8],[9],[10],[11]. Studies have been conducted in Gumi, South Korea, by Lee and Kim [12]. Another study based in Yeongnam, South Korea, was conducted to make day-ahead predictions [13]. They have used meteorological data from the Korea Meteorological Administration (KMA). Their Meteorological data did not include high correlation parameters like humidity, temperature, and solar elevation.

Other than these Machine Learning and Deep Learning techniques, various statistical models have also been used.
It is seen that models based on Numerical Weather Prediction (NWP) systems such as European Centre for Medium-Range Weather Forecast (ECMWF) and the Global Forecast System (GFS) are a good alternative for next day prediction of photovoltaic cells [14]. Daily and monthly solar radiation prediction have also been made with the help of statistical models such as ARIMA (Autoregressive Integrated Moving Average) [15]. Hybrid models combining deep learning and statistical methods have been used and have delivered promising results [16]. Usage of solar angles such as Zenith Angle and Azimuth Angle in conjunction with weather data improve the performance of prediction models [17].
.

III. METHODOLOGY

A. Data-set

To obtain accurate machine learning predictions, we need a good quality dataset with appropriate attributes. The dataset which we have chosen is from a 10MW solar power plant of GPCL (Gujarat Power Corporation Limited) spread across 35 acres. The dataset contains the amount of solar power generated in 5 years. The dataset contains daily values of solar power generated starting from 1st May 2015. Feature scaling was also performed on the dataset. Weather is also an important parameter for the prediction of SE output. In order to obtain higher and better accuracy, we have used weather attributes from NASA Power Project [18]. The location of the GPCL power plant was entered, which is situated in Patan District of Gujarat (23°53'57.9 "N, 71°13'21.3" E).

TABLE I. DESCRIPTION OF DATA-SET ATTRIBUTES

S. No.	*Attributes*	*Format*	*Description*
1	Grid Availability	Numeric	Percentage or factor of grid available for energy procurement.
2	Equipment Availability	Numeric	Percentage of time equipment is used for production
3	Solar Irradiation	Decimal	Solar radiant power incident on a surface per unit area (W-h/m2)
4	Output	Numeric	SE in KWh

The attributes for prediction were Grid Availability, Equipment Availability. Apart from these, Solar Irradiation was also used. These factors helped us to determine the output

[2] https://power.larc.nasa.gov/

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:31:34 UTC from IEEE Xplore. Restrictions apply.

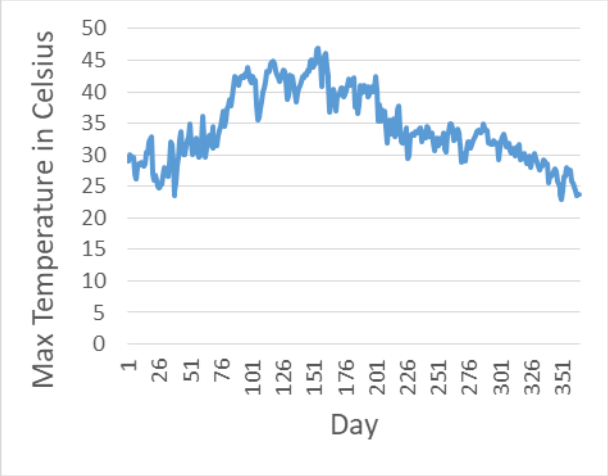


Fig. 1(a) Maximum temperature of each day

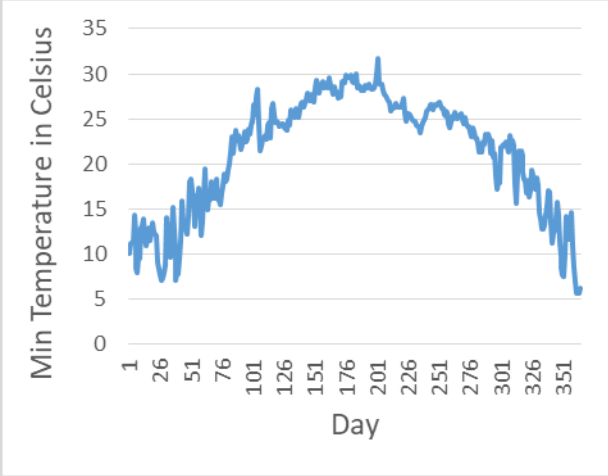


Fig. 1(b) Minimum temperature each day

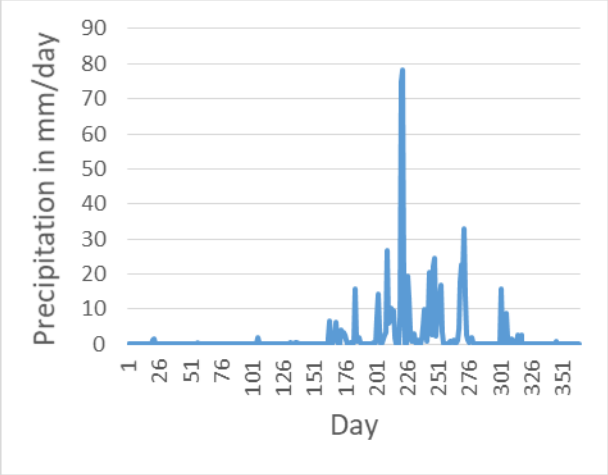


Fig. 1(c) Precipitation of each day

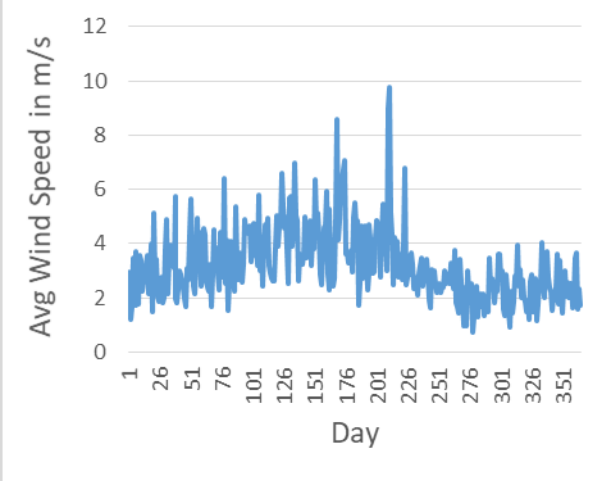


Fig. 1(d) Average Wind Speed of each day

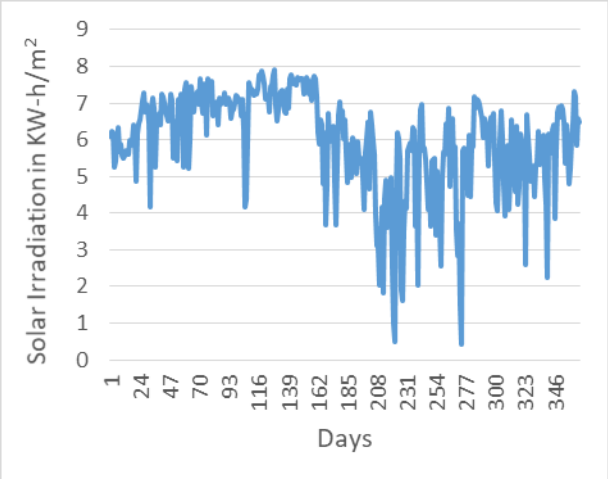


Fig. 1(e) Solar Irradiance of each day

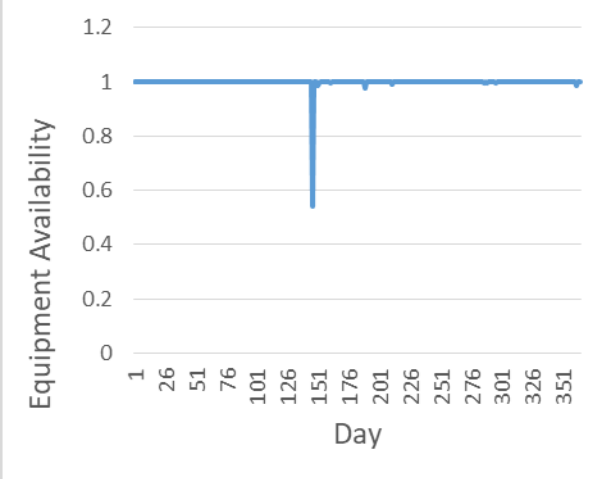


Fig. 1(f) Equipment availability of each day

Fig. 1. Parameter for the Year 2019 (District Patan, Gujarat). X-Label=Day of the Year 2019

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:31:34 UTC from IEEE Xplore. Restrictions apply.

power. The attributes of the dataset, with its data format and their description, are mentioned in Table 1. Fig 1 shows the graph of weather parameters in the year 2019.

B. Model

We have used the Decision Tree Regression Model for our study [19],[20],[21]. Decision Tree Regression is similar to Decision Tree Classifiers, but both serve different purposes, as the name suggests. In this model, we select a variable to split the nodes and evaluate the splits using splitting criteria like Mean Absolute Error (MAE), Mean Square Error (MSE), Poisson function, etc. A tree continues to split until it achieves a minimum sample split parameter which defines the minimum number of samples that are required by the tree to consider splitting a node into 2. The depth of a decision tree must be kept in check as a deeper tree has high variance and low bias while a shorter tree will have a high bias and low variance. Thus, in our study, we have experimented with both the splitting criteria and maximum depth of the tree to obtain the maximum accuracy out of the model, as shown in Table 2.

C. Validation

To ascertain the performance of our model, we have used the Mean Absolute Percentage Error (MAPE) loss function. The mathematical formula for MAPE is given by equation 1. We preferred MAPE as our loss function because we had many outliers in our data-set [22].

We have 56 months of continuous data, and due to its chronological nature, we decided to keep one year of data-set as test data (21.4% of the data-set), as shown in Fig 2. It helped us analyze the performance of our model in all possible weather conditions of Patan District, Gujarat.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Actual\ Value_i - Prediction\ Value_i}{Actual\ Value_i}\right| \quad (1)$$

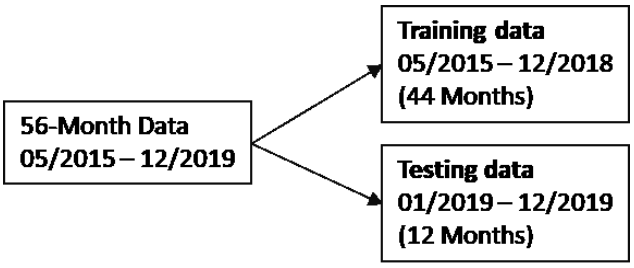


Fig. 2. Dataset split

IV. Results

A. Model Comparison

We have split our data-set into training and testing data-set. Our test data predicts solar power production (KWh) in the year 2019. Then these predictions will be compared to the actual output power in 2019. The result and their error are shown in Table 2 and Fig 3.

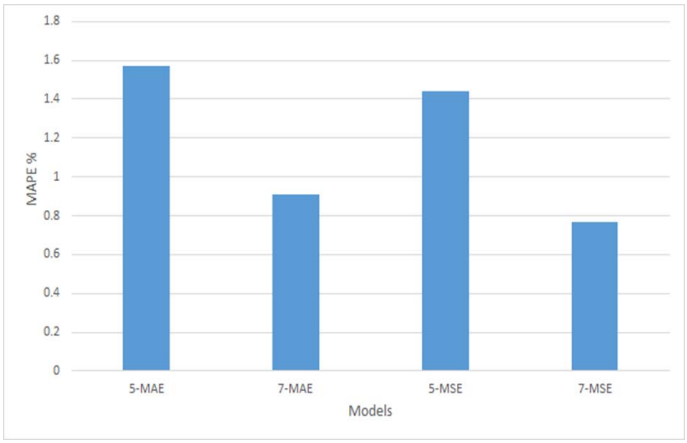


Fig. 3. Bar Graph to Compare Different Models

The Fig 3 is the bar graph of Mean Absolute Percentage Error (MAPE) on the y-axis with the different parameters of the Decision Tree Regressor model on the x-axis. We have four different models for comparison. The model labeled as 5-MAE is the Decision Tree Regressor of maximum depth five and Mean Absolute Error (MAE) with its splitting criterion. Similarly, the model labeled as 7-MSE is a Decision Tree of maximum depth seven and Mean Squared Error (MSE) as its splitting criterion. Similarly, we have labeled 7-MAE and 5-MSE in the graph. The table below shows the values of MAPE for different models.

TABLE II. Splitting Criteria of Decision Tree with its Error Percentage

S. No.	*Splitting Criteria*	*Max Depth*	*MAPE*
1	Mean Absolute Error	5	1.57
2	Mean Absolute Error	7	0.91
3	Mean Squared Error	5	1.44
4	Mean Squared Error	7	0.77

The results shown above are obtained after training our model. The model used was Decision Tree Regressor which gave us satisfactory results. We have obtained results from various splitting criteria and max depth of the Decision Tree. The splitting criteria used were MAE and MSE. The **max_depth** parameter of the Decision Tree was kept 5 and 7. After setting the above parameters, we found the difference between predicted and actual output power production in KWh. The MAPE is computed for each parameter to show deviation from the actual output. We found the least error when the **max_depth** of the tree was seven, and the splitting criteria of the Decision Tree Regressor was MSE. Fig 4 shows

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:31:34 UTC from IEEE Xplore. Restrictions apply.

the graph between actual and predicted power generated (KWh) for all four parameters in the year 2019.

B. *Weather Based Observation*

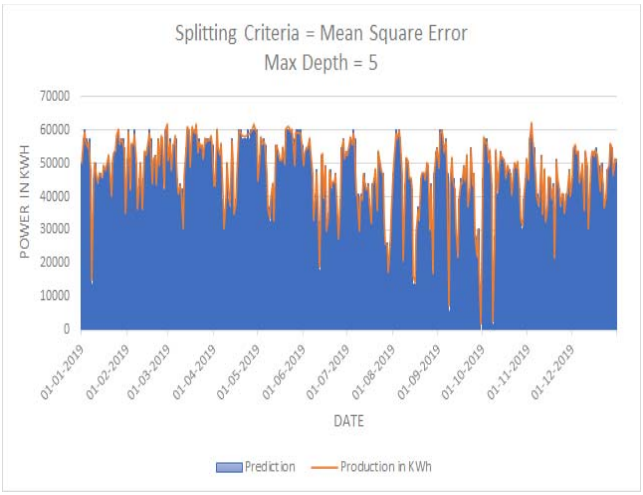


Fig. 4(a) MSE with max depth = 5

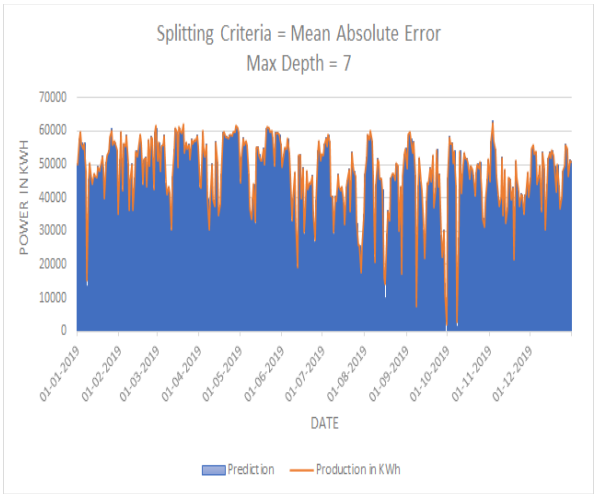


Fig. 4(b) MAE with max depth = 7

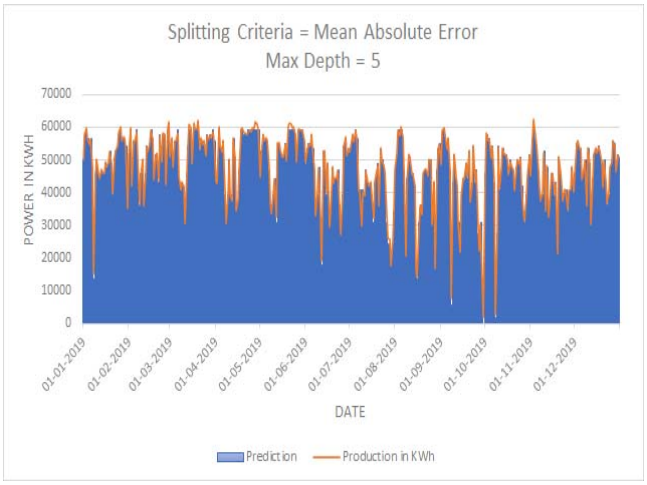


Fig. 4(c) MAE with max depth = 5

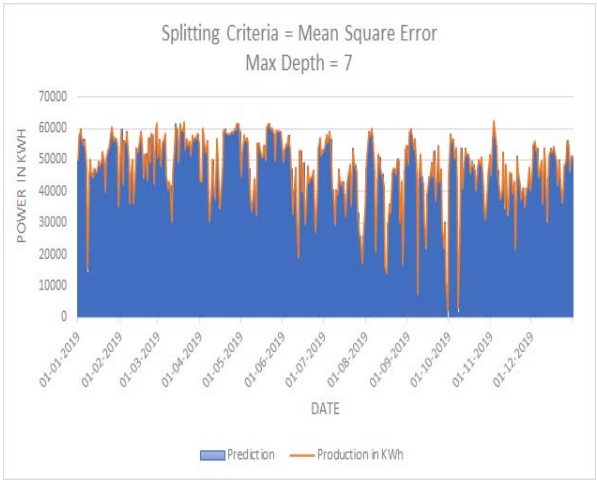


Fig. 4(d) MSE with max depth = 7

Fig. 4. Graph between Actual and Predicted Power Generation for 2019(test data)

The output power production of a SE plant highly depends on the weather conditions of the area. In Fig 5, we have plotted Mean Absolute Percentage Error (MAPE) against all the months of 2019(our test data) using the parameters, maximum depth as seven, and the splitting criteria of the Decision Tree Regressor as Mean Squared Error (MSE)**.** We can observe that the MAPE for the months of September and October is above the MAPE of 2019, the reason being higher precipitation in these two months. During these two months, the Patan district received high rainfall, which increased the average precipitation. Also, the average wind speed was very low for this period. Due to this, solar irradiance was also impacted. It resulted in low power generation compared to other months of the year. High solar irradiation and low average precipitation were observed in the months of February and March 2019. These conditions are highly favorable for a solar power plant, and the error percentage was comparatively lesser for these months as the weather was largely stable.

V. CONCLUSION

SE has been one of the fastest growing renewable energy in recent times, and with the cost coming down to just Rs.36/W for 100KW to 500KW systems[3], we see no indications of a slowdown in the near future. The main challenges that one encounters while planning for a solar-based power generation system are firstly Storage of Energy as energy generation only takes place when solar radiations are present and secondly, the weather-dependent nature of this

[3] https://mercomindia.com/mnre-benchmark-costs-rooftop-solar/

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:31:34 UTC from IEEE Xplore. Restrictions apply.

arrangement. The first challenge will eventually be solved when the battery technology will evolve, especially with the Booming Electric Vehicle industry and the second challenge is what we aim to solve using this study. An accurate prediction ahead in time can help us save carbon emissions as one can reduce dependency on other fossil fuel-based sources of energy and also save up economically. This study on 56 months of data collected from a 10MW installation in Patan District, Gujarat, India, has yielded satisfactory results. This study demonstrates that we can develop a stable grid even with renewable sources of energy which currently is not possible.

VI. Future Works

In the future, it will be fascinating to expand this research further. Studies can be done on various solar power plants to predict their power generation. Similar studies can be performed for different parts of India where there are variations in climatic conditions because the demand for cleaner sources of energy is increasing every day [23],[24]. Several other Artificial Intelligence techniques can be applied to predict the production of Photovoltaic cells like Reinforcement Learning, Deep Learning, etc. We can also use these prediction methods to forecast the power generated through other renewable sources of energy since other sources of energy could also be available in abundance due to differences in geographical conditions.

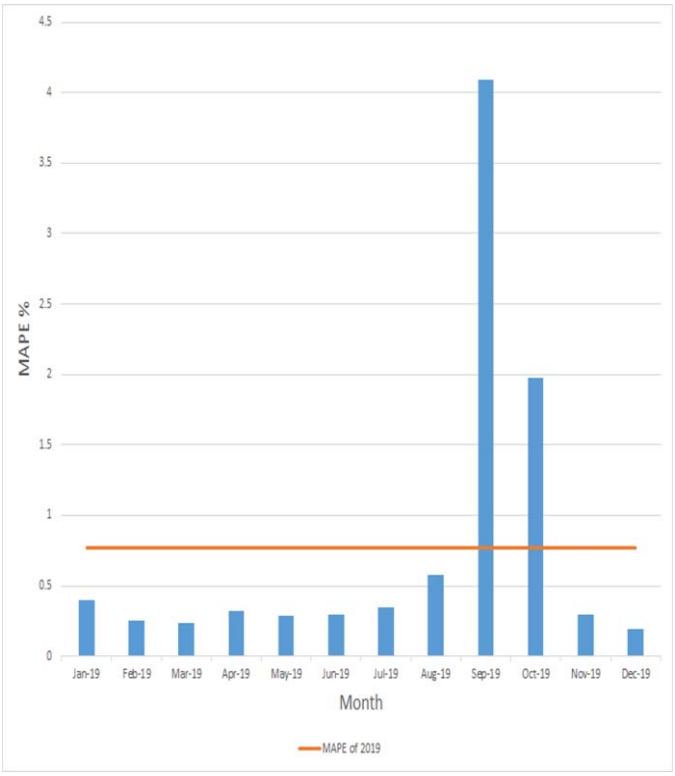


Fig. 5. MAPE for each month of 2019 using model with Model 4 in Table II

References

[1] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de-Pison, and F. Antonanzas-Torres, "Review of photovoltaic power forecasting," *Solar Energy*, vol. 136. pp. 78–111, 2016, doi: 10.1016/j.solener.2016.06.069.

[2] C. Voyant *et al.*, "Machine learning methods for solar radiation forecasting: A review," *Renewable Energy*, vol. 105. pp. 569–582, 2017, doi: 10.1016/j.renene.2016.12.095.

[3] A. Torres-Barrán, Á. Alonso, and J. R. Dorronsoro, "Regression tree ensembles for wind energy and solar radiation prediction," *Neurocomputing*, vol. 326–327, pp. 151–160, 2019, doi: 10.1016/j.neucom.2017.05.104.

[4] A. Khandakar *et al.*, "Machine Learning Based Photovoltaics (PV) Power Prediction Using Different Environmental Parameters of Qatar," *Energies*, vol. 12, no. 14, p. 2782, 2019, doi: 10.3390/en12142782.

[5] M. Aslam, J.-M. Lee, H.-S. Kim, S.-J. Lee, and S. Hong, "Deep Learning Models for Long-Term Solar Radiation Forecasting Considering Microgrid Installation: A Comparative Study," *Energies*, vol. 13, no. 1, p. 147, 2019, doi: 10.3390/en13010147.

[6] Y. Sun, V. Venugopal, and A. R. Brandt, "Short-term solar power forecast with deep learning: Exploring optimal input and output configuration," *Sol. Energy*, vol. 188, pp. 730–741, 2019, doi: 10.1016/j.solener.2019.06.041.

[7] F. Wang, Z. Zhen, B. Wang, and Z. Mi, "Comparative Study on KNN and SVM Based Weather Classification Models for Day Ahead Short Term Solar PV Power Forecasting," *Appl. Sci.*, vol. 8, no. 1, p. 28, 2017, doi: 10.3390/app8010028.

[8] M. W. Ahmad, J. Reynolds, and Y. Rezgui, "Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees," *J. Clean. Prod.*, vol. 203, pp. 810–821, 2018, doi: 10.1016/j.jclepro.2018.08.207.

[9] C. Persson, P. Bacher, T. Shiga, and H. Madsen, "Multi-site solar power forecasting using gradient boosted regression trees," *Sol. Energy*, vol. 150, pp. 423–436, 2017, doi: 10.1016/j.solener.2017.04.066.

[10] J. Fan *et al.*, "Comparison of Support Vector Machine and Extreme Gradient Boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in China," *Energy Convers. Manag.*, vol. 164, pp. 102–111, 2018, doi: 10.1016/j.enconman.2018.02.087.

[11] A. Ahmed Mohammed and Z. Aung, "Ensemble Learning Approach for Probabilistic Forecasting of Solar Power Generation," *Energies*, vol. 9, no. 12, p. 1017, 2016, doi: 10.3390/en9121017.

[12] D. Lee and K. Kim, "Recurrent Neural Network-Based Hourly Prediction of Photovoltaic Power Output Using Meteorological Information," *Energies*, vol. 12, no. 2, p. 215, 2019, doi: 10.3390/en12020215.

[13] S.-G. Kim, J.-Y. Jung, and M. Sim, "A Two-Step Approach to Solar Power Generation Prediction Based on Weather Data Using Machine Learning," *Sustainability*, vol. 11, no. 5, p. 1501, 2019, doi: 10.3390/su11051501.

[14] R. Aler, R. Martín, J. M. Valls, and I. M. Galván, "A study of machine learning techniques for daily solar energy forecasting using numerical weather models," *Stud. Comput. Intell.*, vol. 570, pp. 269–278, 2015, doi: 10.1007/978-3-319-10422-5_29.

[15] M. Alsharif, M. Younes, and J. Kim, "Time Series ARIMA Model for Prediction of Daily and Monthly Average Global Solar Radiation: The Case Study of Seoul, South Korea," *Symmetry (Basel).*, vol. 11, no. 2, p. 240, 2019, doi: 10.3390/sym11020240.

[16] M. AlKandari and I. Ahmad, "Solar power generation forecasting using ensemble approach based on deep learning and statistical methods," *Applied Computing and Informatics*. 2019, doi: 10.1016/j.aci.2019.11.002.

[17] F. Jawaid and K. Nazirjunejo, "Predicting daily mean solar power using machine learning regression techniques," in *2016 6th International Conference on Innovative Computing Technology, INTECH 2016*, 2017, pp. 355–360, doi: 10.1109/INTECH.2016.7845051.

[18] A. Sparks, "nasapower: A NASA POWER Global Meteorology,

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:31:34 UTC from IEEE Xplore. Restrictions apply.

Surface Solar Energy and Climatology Data Client for R," *J. Open Source Softw.*, vol. 3, no. 30, p. 1035, 2018, doi: 10.21105/joss.01035.

[19] F. Mola, "Classification and Regression Trees Software and New Developments," 1998, pp. 311–318.

[20] T. Hastie, R. Tibshirani, and J. Friedman, "Linear Methods for Regression," 2009, pp. 43–99.

[21] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[22] B. Jiang and C. Ding, "Revisiting L2,1-Norm Robustness with Vector Outlier Regularization," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 31, no. 12, pp. 5624–5629, Dec. 2020, doi: 10.1109/TNNLS.2020.2964297.

[23] D. S. V., "Data Mining based Prediction of Demand in Indian Market for Refurbished Electronics," *J. Soft Comput. Paradig.*, vol. 2, no. 2, pp. 101–110, 2020, doi: 10.36548/jscp.2020.2.007.

[24] A. Chandy, "SMART RESOURCE USAGE PREDICTION USING CLOUD COMPUTING FOR MASSIVE DATA PROCESSING SYSTEMS," *J. Inf. Technol. Digit. World*, vol. 01, no. 02, pp. 108–118, 2019, doi: 10.36548/jitdw.2019.2.006.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:31:34 UTC from IEEE Xplore. Restrictions apply.

Materials Chemistry and Physics 270 (2021) 124856

Contents lists available at ScienceDirect

Materials Chemistry and Physics

journal homepage: www.elsevier.com/locate/matchemphys

Strong enhancement in structural, dielectric, impedance and magnetoelectric properties of $NdMnO_3$ - $BaTiO_3$ multiferroic composites

S. Shankar [a,b], O.P. Thakur [c,*], M. Jayasimhadri [b,**]

[a] Experimental Research Laboratory, Department of Physics, ARSD College, University of Delhi, 110021, India
[b] Department of Applied Physics, Delhi Technological University, 110042, India
[c] Materials Analysis and Research Laboratory, Department of Physics, NSUT, New Delhi, 110078, India

HIGHLIGHTS

- (1-x)$NdMnO_3$ –(x)$BaTiO_3$ multiferroic composites have been synthesized via solid state reaction method.
- Significant improvement is observed in dielectric, impedance and ferroelectric properties of $NdMnO_3$ with BT addition.
- Multiferroism and magnetoelectricity displayed by $NdMnO_3$ -$BaTiO_3$ composites.
- These results are related to lattice stability and thermally activated hopping mechanism.

ARTICLE INFO

Keywords:
Composites
X-ray diffraction
Dielectric properties
Multiferroics

ABSTRACT

Multiferroic composites are advantageous from technological perspective because of their flexible electrical and magnetic properties. The multiferroic composites of (1-x)$NdMnO_3$ – (x)$BaTiO_3$ (x = 0, 0.1 and 0.2) were synthesized via solid state reaction method. The structural analysis from XRD studies of (1-x)$NdMnO_3$-(x)$BaTiO_3$ exhibits distinct peaks of orthorhombic $NdMnO_3$ (NMO) and tetragonal $BaTiO_3$ (BT) phases with absence of any other impurity phases in the composites. The frequency variation of dielectric measurements reveals step-like decrease and confirm Maxwell-Wagner Polarization associated with improved conducting nature in NMO-BT composites. The NMO-20BT composite exhibited giant dielectric behavior owing from lattice stability and hopping mechanism. The impedance measurements confirm improved resistivity in NMO-BT composites and confirm thermally activated non-Debye type conduction behavior in NMO-BT in them. The occurrence of saturated ferroelectric polarization composites along with enhanced remnant magnetization and improved magnetoelectric coupling in NMO-20BT composite due to lattice stability establishes multiferroism in NMO-BT composites.

1. Introduction

Multiferroic magnetoelectric composite materials have substituted for single-phase multiferroics in the last decades [1,2]. The magnetoelectric coupling in multiferroic composites is enticed through physical strain mediated interaction amongst ferroelectric and ferromagnetic phases and are ascribed as piezo/magneto-striction [3–5]. Large number of composites based on ferrites as well as manganites for magnetic constituent and barium as well as lead based titanates for the ferroelectric phase have been investigated for multiferroics and magnetoelectric properties [6–9].

Rare earth manganites $ReMnO_3$ (Re = Rare earth) are vastly studied systems due to the occurrence of captivating features like magnetic orderings, metal-insulator transition, colossal magnetoresistance and multiferroism originating from charge spin and lattice dynamics [10–12]. $ReMnO_3$ finds their application in information storage devices owing from their broad antiferromagnetism Neel temperature range of 620–740 K [13,14]. These manganites exhibit peculiar structural and charge conductivity transport phenomena that are governed by the behavior of transitional metal (Mn) ions with the neighboring (Re) ions. These manganites possess distinctive fascinating electronic configuration due to the hybridization of 3d and 2p orbital of Mn and oxygen,

* Corresponding author.
** Corresponding author.
E-mail addresses: opthakur@yahoo.com (O.P. Thakur), jayaphysics@yahoo.com (M. Jayasimhadri).

https://doi.org/10.1016/j.matchemphys.2021.124856
Received 17 January 2021; Received in revised form 17 May 2021; Accepted 16 June 2021
Available online 20 June 2021
0254-0584/© 2021 Elsevier B.V. All rights reserved.

respectively. Manganites of La, Pr or Nd usually exhibit orthorhombic structure having *Pnma* space group [15,16]. In recent years, the magnetoelectric and magnetodielectric explorations in $ReMnO_3$ have enticed enormous interest due to the presence of different origins of multiferroics properties, the most common being the double exchange interaction (DE) arising from the presence of mixture of Mn^{2+}/Mn^{3+} ions as given below in Equation (1) [17,18].

$$Mn^{2+} + e_g^2 \leftrightarrow Mn^{3+} \quad (1)$$

The hopping of doubly degenerate e_g^2 electron induces the magnetic transition. Amongst the existing mixed valent manganites, $NdMnO_3$ (NMO) is the best possible candidate to exploit for enhancement in ME properties due to the existence of structural disordering originating from tilting of MnO_6 octahedra. This tilting provides an antiferromagnetic A to E − type magnetic transition required for spontaneous multiferroism [11,19]. NMO exhibits insulating ferromagnetic nature which can be tailored by doping Nd^{3+} site or Mn^{3+} site by divalent atoms such as Ba^{2+}, Sr^{2+}, Cd^{2+}, Ca^{2+}, Cu^{2+} and Zn^{2+} ions [10,11,17,20]. The studies on NMO doped systems have shown interesting results, however, the composite study of NMO is barely countable. The studies on NMO based composites show the advantage of composite system instead of parent NMO [21,22].

Barium titanate ($BaTiO_3$) is a conventional ceramic with excellent ferroelectric and dielectric properties [23]. $BaTiO_3$ (BT) is widely used in electronic devices such as photoelectrodes, layered capacitors and sensors [23]. It is one of the best suitable material employed as dielectric in electronic devices and exhibits excellent piezoelectricity and ferroelectricity [24,25].

The composites of bulk ferromagnetic NMO with ferroelectrics are expected to show strong magnetoelectric coupling close to room temperature owing to the product property of multiferroics magnetoelectric composites. In this regard, the bulk composites of $NdMnO_3$ - $BaTiO_3$ will be an interesting study as it will involve the multiferroic magnetoelectric nature of NMO combined to the strong ferroelectric nature of BT. This work describes the synthesis $NdMnO_3$ - $BaTiO_3$ (NMO-BT) composite to study the effect of BT in NMO on the structural, dielectric, magnetic, and ferroelectric *P-E* and magneto-electric properties.

2. Experimental

2.1. Materials and methods

The (1-*x*)NMO-*x*BT composites ($x = 0$, 0.1 and 0.2) were synthesized using solid state reaction route. High purity (≥99%) of Central Drug House (CDH) make of chemicals of Nd_2O_3, MnO_2, $BaCO_3$ and TiO_2 were used as initial ingredients in stoichiometric ratio. The synthesis conditions have been tabulated in Table 1. The initial oxides for $BaTiO_3$ and $NdMnO_3$ respectively were stoichiometrically weighed, mixed and ground for 8 h. The mixed powders of $BaTiO_3$ and $NdMnO_3$ were calcined in muffle furnace for 1000 °C and 900 °C respectively. The calcined powder was ground by mixing stoichiometrically for preparing desired $NdMnO_3$, $0.9NdMnO_3 - 0.1BaTiO_3$ and $0.8NdMnO_3 - 0.2BaTiO_3$ compositions. The composition powders were mixed with poly vinyl alcohol binder and hard pressed into cylindrical pellets of diameter of 10 mm. These pellets were sintered at 950 °C for 10 h. The sintered composites were named as NMO ($NdMnO_3$), NMO-10BT ($0.9NdMnO_3 - 0.1BaTiO_3$) and NMO-20BT ($0.8NdMnO_3 - 0.2BaTiO_3$).

Table 1
Synthesis conditions of $BaTiO_3$, $NdMnO_3$ and $NdMnO_3$ - $BaTiO_3$ composites.

Composition	Calcination temperature (°C)	Duration (h)	Sintering temperature (°C)	Duration (h)
$BaTiO_3$	1000	8	–	–
$NdMnO_3$	900	8	950	10
NMO-10BT	–	–	950	10
NMO-20BT	–	–	950	10

2.2. Characterization

The powders crushed from the sintered pellets were used for recording the XRD patterns, IR spectra, Raman Spectra and VSM. The X-ray diffraction (XRD) studies have been done with a Bruker D8 - Advance X-ray Diffractometer employing Cu-K_α radiations and wavelength of 1.5418 Å. The Fourier transform infrared spectroscopy (FT-IR) measurements were made using Bruker FT-IR instrument operating in the wave number range of 400–1000 cm^{-1}. The Raman spectra have been observed employing Renishaw Raman Spectrometer. The room temperature magnetic hysteresis *M-H* measurements were recorded using vibrating sample magnetometer (VSM). The sintered pellets were silver coated on surface for making better contact for dielectric properties. The dielectric parameters were measured using Waynekerr 6500 P high frequency LCR meter setup over the frequency range of 100 Hz −1 MHz and at fixed temperatures of 100, 200, 300 and 400 °C respectively. The ferroelectric measurements were made from Marine India *P-E* hysteresis setup at a frequency of 50 Hz, electric field of up to 4 kV/cm. The evidence of magnetoelectric coupling was made by studying the ferroelectric measurements for magnetic field varying from 0 to 5000 Oersted.

3. Results and discussion

3.1. Structural measurements

XRD spectra of NMO, BT and NMO-BT composites have been recorded at room temperature and plotted in Fig. 1. The XRD patterns have been indexed the perovskite $NdMnO_3$ and $BaTiO_3$ phase with JCPDS # 25–1149 for orthorhombic phase and JCPDS # 05–0626 for tetragonal phase, respectively. The XRD patterns of NMO and BT samples show pure decomposed single phase with no presence of any Nd, Ba, Mn and Ti oxides and indicates high phase purity. The sharpened peaks in the XRD spectra are owing to the strong crystallinity in NMO and BT compounds. The XRD patterns of NMO-10BT and NMO-20BT show the evolution and co-existence of both individual phase of NMO and BT and absence of any other third phase, which confirms the formation of high purity NMO-BT composites.

The impact of BT addition in NMO is understood by using the Debye Scherrer Equation (2) and Williamson-Hall (W–H) Equation (3) [26–28] given as

$$D = \frac{k\lambda}{\beta Cos\theta} \quad (2)$$

$$\frac{\beta Cos\theta}{\lambda} = \frac{1}{\varepsilon} + \frac{\eta Cos\theta}{\lambda} \quad (3)$$

where $k = 0.89$, λ is the wavelength of X-rays used, β is full width at half maximum (FWHM) of the diffraction peak, θ is angle of diffraction, ε denote effective crystallite size and η denote effective strain. The W–H equation is used to obtain accuracy in crystallite size and strain. Fig. 2 (a) shows the W–H plot for BT, NMO and NMO-BT composites. There is a decreasing trend in the strain obtained from W–H plot indicating that the incorporation of BT in NMO lattice has led to stabilization. The crystallite size obtained from Scherrer equation is plotted in Fig. 2 (b). The crystallite size (d_{NMO}) in NMO is found to increase in NMO-BT composite as compared to NMO and similarly the crystallite size (d_{BT}) in BT is found to decrease suggesting a lattice distortion. Moreover, the crystallite size obtained from W–H plot is in agreement with the crystallite size obtained using Scherrer equation.

The crystal lattice distortion of NMO due to BT can be predicted on the basis of tolerance factor formula of ABO_3 type perovskites expressed as Equation (4) [29].

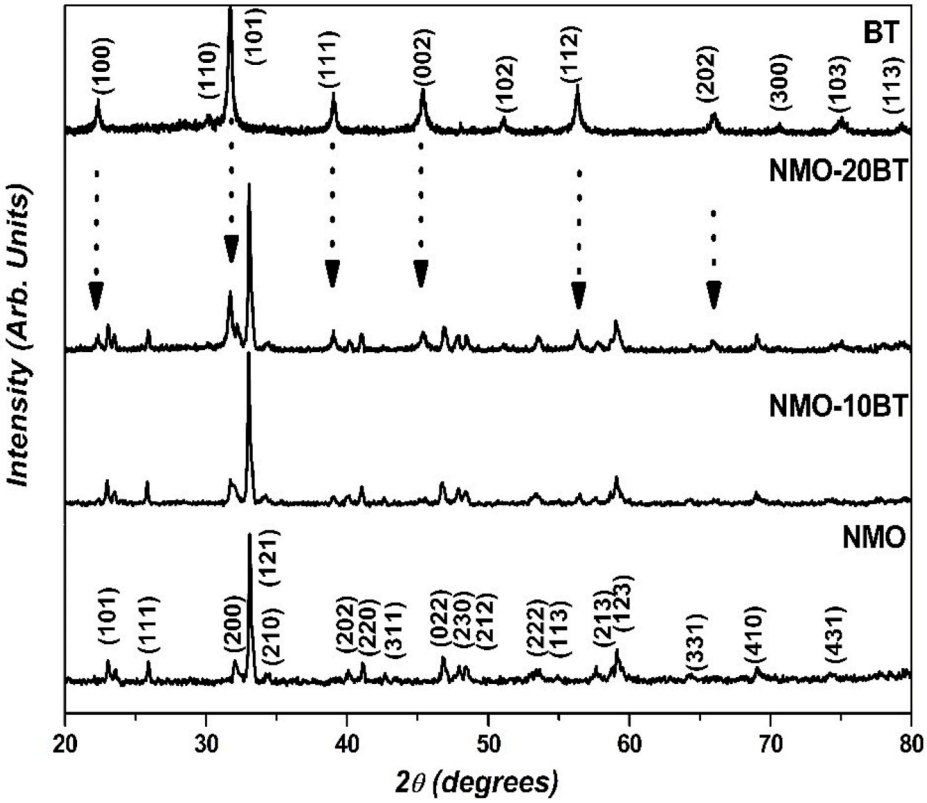


Fig. 1. XRD patterns of NMO, BT and NMO-BT composites.

$$\gamma = \frac{r_A + r_o}{\sqrt{2}(r_B + r_o)} \tag{4}$$

where r_i denote ionic radius of atoms i = A, B and O respectively (in Angstroms). The calculated tolerance factor based on the standard ionic radii is tabulated in Table 2 along with corresponding ionic radii used. It can be seen from the table that tolerance factor of NMO and BT lie in their respective range of orthorhombic ($0.75 < \gamma < 0.9$) and tetragonal ($\gamma > 1$) phases respectively. For NMO-10BT composite, the predicted tolerance factor corresponds to 1.014 and indicate a tetragonal structure. The structural modification from orthorhombic structure of NMO to predicted tetragonal structure in NMO-10BT composite could be due to infusion effect of tetragonal lattice of BT in orthorhombic lattice of NMO. This infusion could be related to observed decrease in computed effective strain and suggest a lattice rearrangement of atoms. Similarly, the predicted tolerance factor for NMO-20BT composite decreases to 0.84 corresponding to orthorhombic structure. This suggests that further infusion of tetragonal lattice of BT in orthorhombic lattice of NMO is not impacting the structure modification. In other words, the lattice is stabilized in terms of symmetry and atomic arrangement on account of observed decrease in effective strain.

FT-IR spectrum is useful in providing information about the behavior of metal ions in the perovskite lattice. The FT-IR spectra of NMO and NMO-BT composites in the wavenumber range of 400–1000 cm^{-1} are plotted in Fig. 3. The orthorhombic phase of manganites as described by factor group analysis consists of the following optical modes: $\Gamma = 7A_{sym}(R) + 7B_{1sym}(R) + 5B_{2sym}(R) + 5B_{3sym}(R) + 8A_{unsym}(IA) + 7B_{1unsym}(IR) + 8B_{2unsym}(IR) + 9B_{3unsym}(IR)$, where 'sym' and 'unsym' denotes gerade and ungerade optical modes. The FT-IR spectra of NMO show minimum or no splitting indicating high ordering in crystal lattice. The presence of strong peak at 556 cm^{-1} relating to non-symmetric vibrations of Mn–O–Mn bond generating from MnO_6 octahedra and broadening in region from 405 to 481 cm^{-1} relating to vibrations of Mn–O bond confirm the formation of manganite NMO [30]. The broad region is originated due to the John-Teller interactions which widens equatorial bond and shortens axial Mn–O bonds. Fig. 3 also shows the FT-IR spectra of NMO-10BT and NMO-20BT. It can be observed that the intense peak of 556 cm^{-1} increases to 562 cm^{-1} in NMO-10BT and shifts to 557 cm^{-1} in NMO-20BT. This shift in intense band located near 552 cm^{-1} relating to vibrations of Ti–O bond arising from TiO_6 octahedra confirms the formation of ferroelectric BT. Moreover, the broadening of this peak reduces with BT incorporation along with increased splitting in the widened region suggesting that Ti–O bonds of BT introduce additional vibrations in the NMO lattice. The reduction in broadening is linked to decrease in localized modes and charge carriers of Ti–O bonds hindering the optical modes of Mn–O–Mn bonds [31]. This softening of FT-IR spectra and increased splitting confirms the formation of NMO-BT composites as validated by XRD studies.

The room temperature Raman spectra of NMO and NMO-BT composites is shown in Fig. 4. The spectra show intense peak at 636 cm^{-1} of B_{1sym} mode relating to O–O vibrations originating from John-Teller interactions and peak at 288 cm^{-1} of A_{sym} mode relating to tilting of MnO_6 octahedra. However, the usual peak at 480 cm^{-1} relating to second order non-symmetric stretching of O–O is absent. The Raman spectra of NMO-BT composites show an improvement in the peaks around 180, 260 and 512 cm^{-1} indicated by '*' relating to the transverse non-symmetry vibration Ti–O (A_T) of phonons and indicate infusion of tetragonal $BaTiO_3$ into orthorhombic $NdMnO_3$ lattice [32] corroborating with XRD and FT-IR studies.

3.2. Dielectric measurements

The dielectric properties of the composites are analyzed with the content of BT in NMO-BT composites. The frequency dependence plots of dielectric constant (ε') for NMO-BT composites is shown in Fig. 5. The NMO-BT composite samples show strong step like decrease above frequency of 10 kHz behavior indicating the presence of dipolar relaxation processes at all temperatures [30]. In the low frequency region,

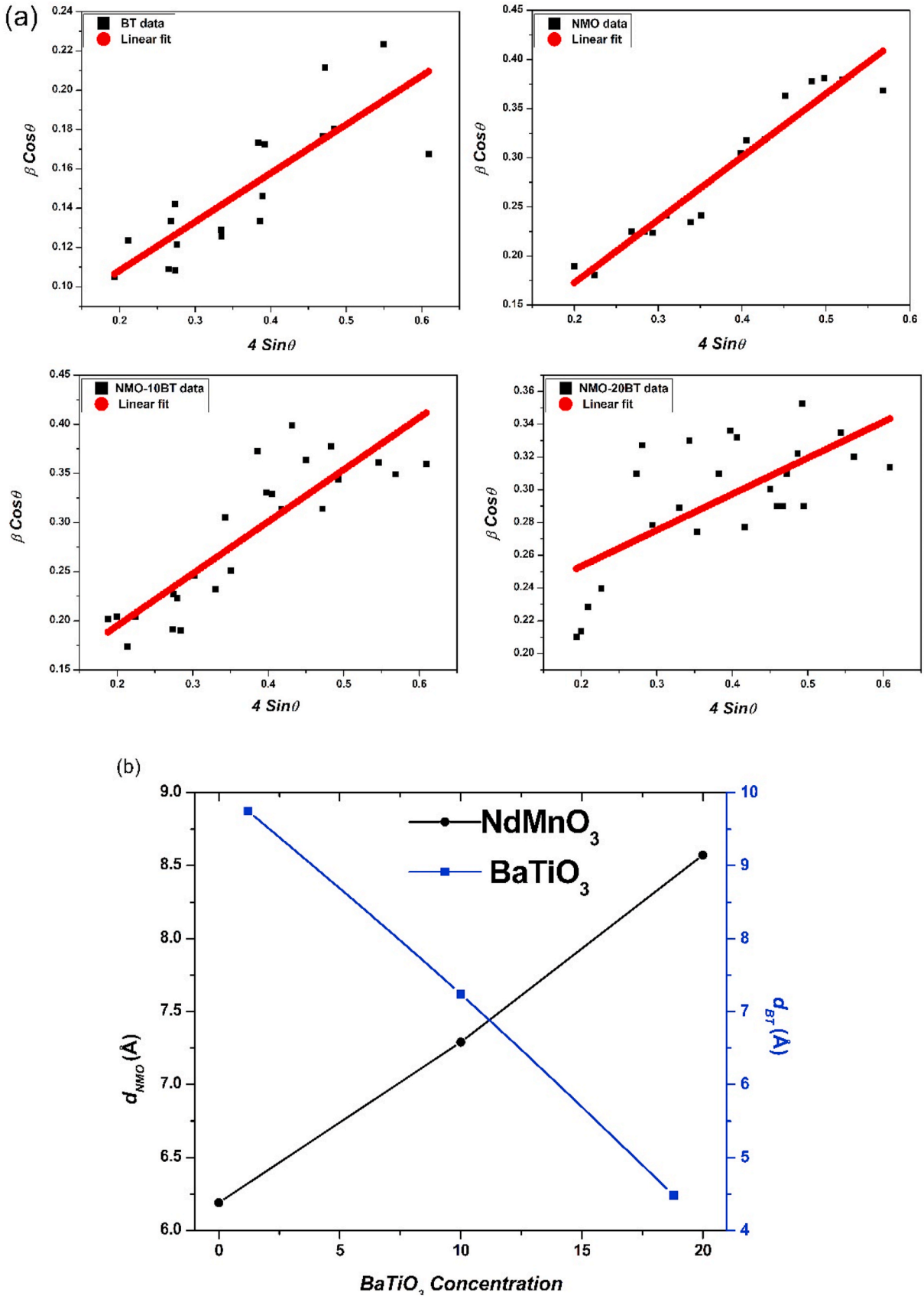


Fig. 2. (a) W–H plots of BT, NMO and NMO-BT composites (b) Variation of crystallite size of NMO and BT in NMO-BT composites.

Table 2
Variation of tolerance factor with BT concentration.

Element	Ionic radii (Å)	Composition	Tolerance Factor γ
Nd	1.16		
Mn	0.83	BT	1.06
Ba	1.38	NMO	0.81
Ti	0.605	NMO-10BT	1.014
O	1.42	NMO-20BT	0.84

dielectric constant is more or less static as electric dipoles are following the oscillations of applied electric field. This behavior is influenced by fundamental conduction mechanism in NMO and NMO-BT composites. Furthermore, at higher frequencies, the dipoles are unable to follow the oscillations of electric field. This dispersion dielectric characteristic attributing to non-Debye relaxations can be modelled as in Equation (5)

$$\varepsilon_r = \varepsilon_{r\infty} + \frac{\varepsilon_{rs} - \varepsilon_{r\infty}}{1 + (i\omega\tau)^{1-\alpha}} \quad (5)$$

where α, τ, $\varepsilon_{r\infty}$ and $\varepsilon_{r\infty}$ denotes constant, relaxation time, dielectric constant at very high frequencies and dielectric constant at static field, respectively [33]. This dispersion in dielectric behavior at higher frequencies is ascribed to relaxor behavior [27]. The enhancement in dielectric permittivity at higher temperatures is attributed to increased conduction in the NMO-BT composites. It can also be observed that the value of dielectric constant at any measured temperature is more in NMO-BT composite than NMO due to the strong dielectric nature. The giant increase in dielectric constant of NMO-20BT composite is related to stabilization in the lattice, as also confirmed by XRD measurements. The probable cause of enhanced dielectric constant in NMO-20BT composite, in terms of charge carriers, could be due to the presence of hopping conductivity phenomena [34].

3.3. Impedance measurements

The electrical properties in composite materials viz., role of grains and grain boundaries can be understood by examining the deviation in the response of impedance as function of applied oscillating field [35]. The real (Z') and imaginary (Z'') components of the complex impedance $Z^* = Z'\text{-}iZ''$, are plotted as Cole-Cole plots for NMO and NMO-BT composites and are displayed in Fig. 6. The effective contributions of grains and grain boundaries in NMO-BT composites can be distinguished and pictured by fitting the Cole-Cole plot to an equivalent circuit consisting of resistors and capacitor as in Equation (6)

$$Z^* = \frac{R_1}{1 + i\omega R_1 C_1} + \frac{R_2}{1 + i\omega R_2 C_2} \quad (6)$$

where ω is the angular frequency of the oscillating field. In the present work, all the samples show non-Debye nature and thus capacitor is substituted by constant phase element (CPE) [33]. The CPE impedance can be represented as $Z^*_{CPE} = \frac{1}{C_c(i\omega)^n}$, where C_c is the capacitance of CPE and n is the power exponent. The Cole-Cole plots of NMO and NMO-BT composites show partial as well as complete semi-circular arcs with their centers lying below the Z' axis. The display of lower curvature confirms the presence of different relaxation processes associated with non-ideal electric behavior in NMO and NMO-BT composites. In addition, the non-Debye behavior could be due to the presence of defects and vacancies in NMO-BT composites. All samples show a decrease in the radius of Cole-Cole plots with increase in temperature from 100 °C to 400 °C and indicate the presence of thermally dominated conduction mechanism. It is also interesting to note that the Cole-Cole plot of NMO-20BT alone doesn't start from origin indicating the improved and strong dielectric nature in the composite [36]. The experimental data has been fitted with R-CPE elements using Equation (6) and the

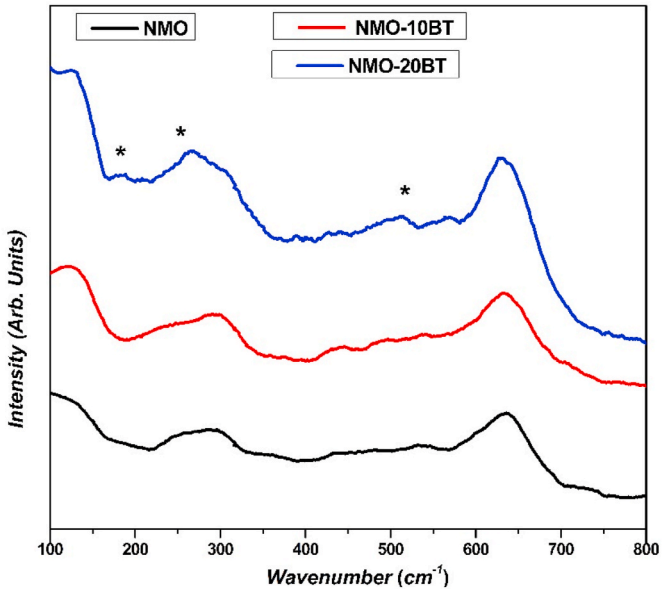


Fig. 4. Raman spectra of NMO-BT composites.

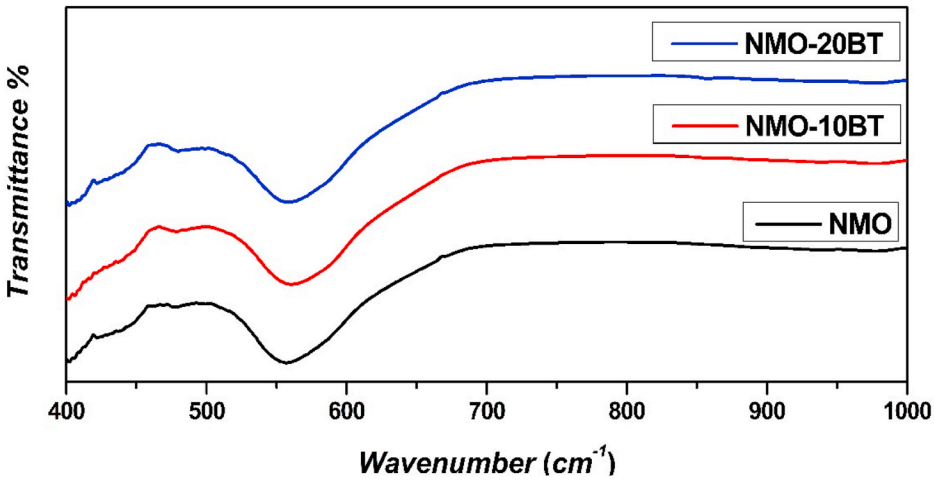


Fig. 3. FT-IR spectra of NMO and NMO-BT composites.

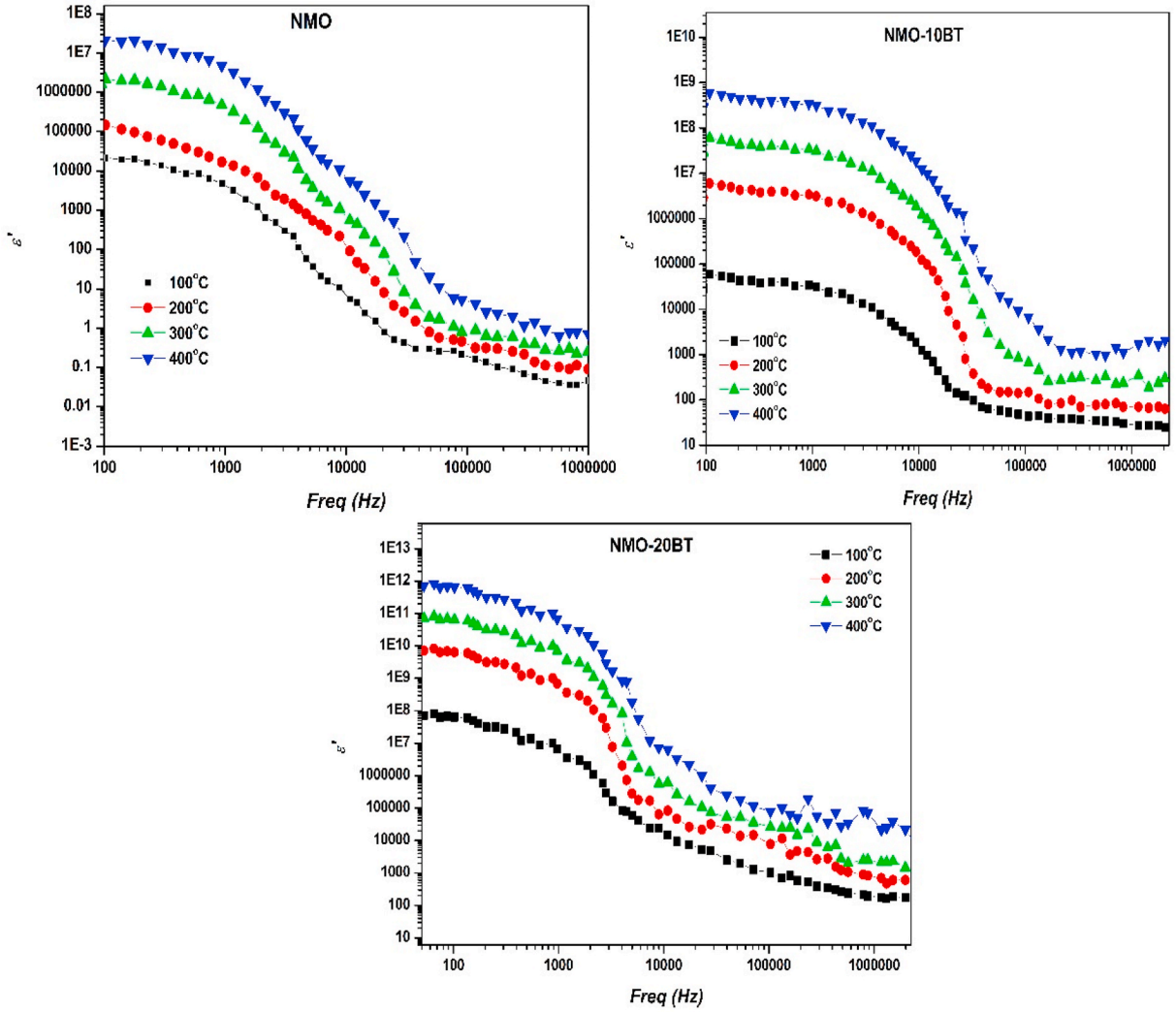


Fig. 5. Frequency dependence of dielectric constant of NMO and NMO-BT composites.

corresponding values can be construed. The experimental data along with equivalent circuit and fitted spectra are shown in Fig. 6. The measured data and R-CPE equivalent fitted spectra show good agreement with each other. The resistance of grains is always greater than that of grain boundaries at all temperatures. The estimated parameters are tabulated in Table 3. The activation energy obtained from the thermal variation of resistance of grains is of the order of 0.46 eV for NMO sample [1].

3.4. Conductivity measurements

The transport behavior of charge carriers can be understood by studying the frequency dispersion of conductivity (σ′) for NMO-20BT composite using Equation (7)

$$\sigma' = \sigma_{dc} + \sigma_{ac} = \sigma_{dc} + A\omega^{s} \quad (7)$$

where constants A and s have usual meanings [37]. Fig. 7 shows static conductivity at all temperatures up to log $\omega \approx 5$ rad s^{-1} and indicates dc conductivity (σ_{dc}). This dc behavior is correlated to the movement of charges across grain boundaries. The low value of conductivity in this region could be due to charge accumulation at grain boundary/electrode interface. On the contrary, the conductivity in Fig. 7 shows strong dispersion and rapidly increases in the high frequency region indicating intensified mobility of multivalent charge carriers of Mn and Ti atoms. The increase in temperature decreases the conductivity suggesting temperature dependent conductivity phenomena of charge carriers in NMO-20BT composite.

The temperature dependence of dc conductivity for NMO-BT is shown in Fig. 8 and the activation energy for NMO and NMO-BT composites obtained using the Arrhenius law as in Equation (8)

$$\sigma_{dc} = \sigma_{o} e^{-E_a/k_B T} \quad (8)$$

where σ_o, k_B, T have usual notations [37]. The activation energies of NMO and NMO-BT composites are less than unity and are mainly due to the presence of single oxidized ions from Mn and Ti atoms produced during synthesis as given by Kroger-Vink defect Equation (9)

$$O_o^x \rightarrow \frac{1}{2}O_2 + V_o^{**} + e^- \quad (9)$$

where V_o^{**} is the ionized metal [38]. The activation energy obtained using thermal variation of resistance of grains is in agreement with the activation obtained from Arrhenius law.

The thermal variation of exponent s for NMO and NMO-BT composites is shown in Fig. 9. Based on the types of conduction mechanism [37], it can be observed and understood that the composites display correlated barrier hopping between the neighboring sites through thermal activation. This hopping behavior is temperature dependent and can be visualized as the movement of multivalent carriers like

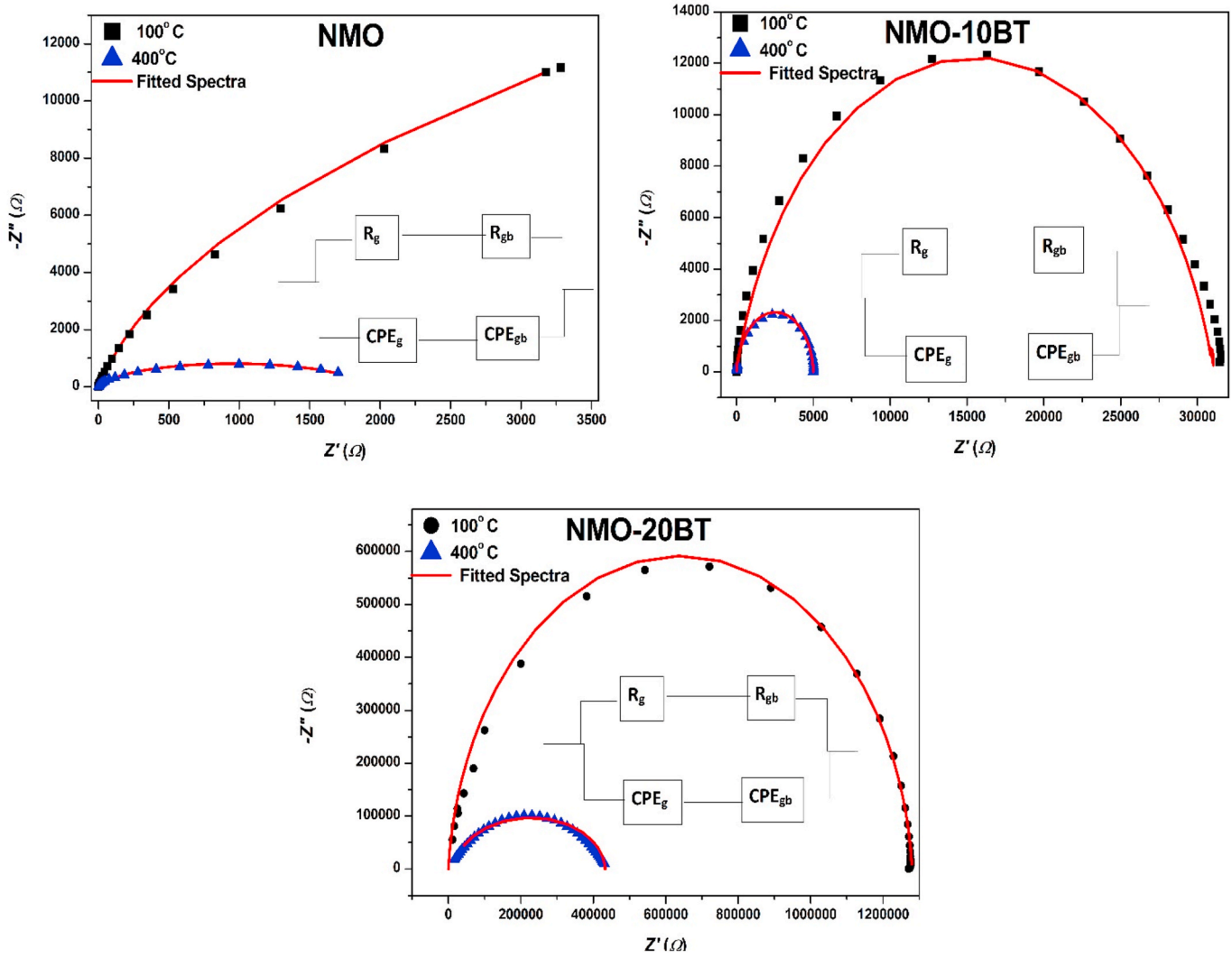


Fig. 6. Cole-Cole plots and equivalent circuit with fitted spectra of NMO-BT composites.

Table 3
Fitted parameters for Cole–Cole plots of NMO-BT composites.

Temp (°C)	*R_g (Ω)*	*CPE_g*	*n_g*	*R_{gb} (Ω)*	*CPE_{gb}*	*^{n}gb*
NMO						
100	2100000	1.1×10^{-9}	0.81	80000	8×10^{-9}	0.65
400	11625	5×10^{-8}	0.67	1150	5×10^{-7}	0.4
NMO-10BT						
100	3300000	2×10^{-9}	0.85	570000	9×10^{-9}	0.73
400	8000	8×10^{-8}	0.7	2796	3×10^{-7}	0.55
NMO-20BT						
100	435000	3×10^{-9}	0.89	6400	2×10^{-8}	0.85
400	4050	4×10^{-8}	0.73	190	8×10^{-7}	0.64

Mn^{3+}/Mn^{2+} ions between nearest neighboring localized sites via hopping mechanism. The values of *s* for NMO and NMO-BT composites remains less than unity and indicates translational motion of charge carriers [39].

3.5. Magnetic measurements

The room temperature magnetization measurements obtained using vibrating sample magnetometer studies magnetization (*M*) at a maximum field of 15000 Oe are plotted in Fig. 10. The *M-H* plots indicate that magnetization of the $NdMnO_3$-$BaTiO_3$ composites increases linearly with the field and confirms the antiferromagnetic behavior of $NdMnO_3$. The occurrence of magnetization in NMO-BT composites is predominantly due to the antiferromagnetic nature of NMO. The

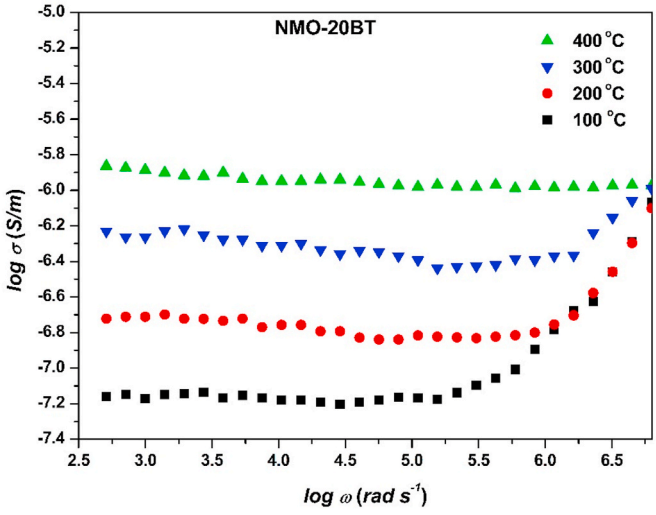


Fig. 7. Variation of log σ′ versus log ω for NMO-20BT composite.

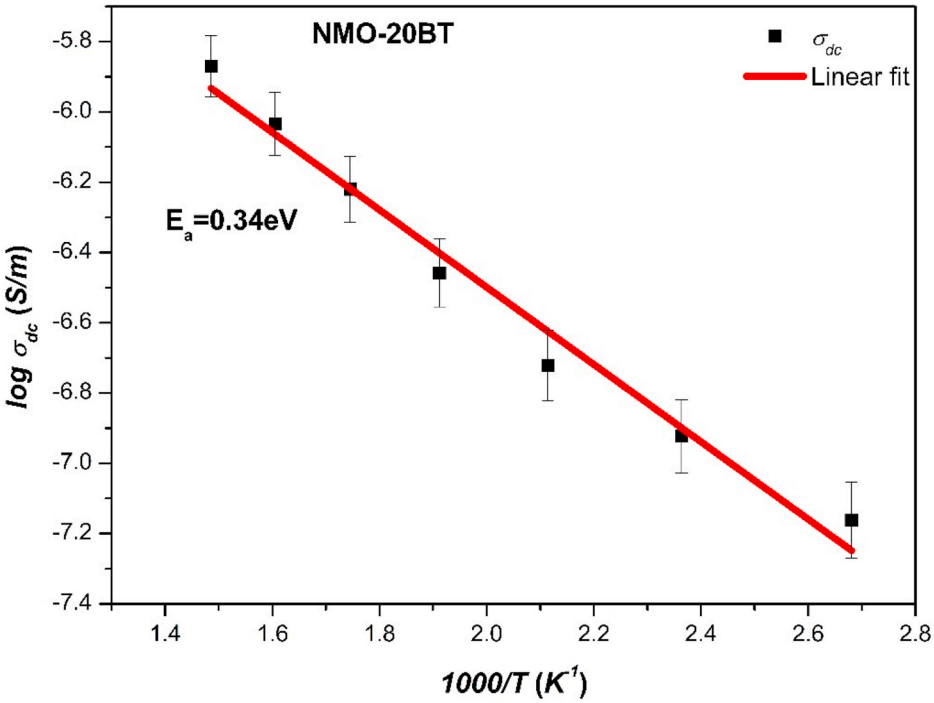


Fig. 8. Plot of DC conductivity versus temperature for NMO-20BT composite.

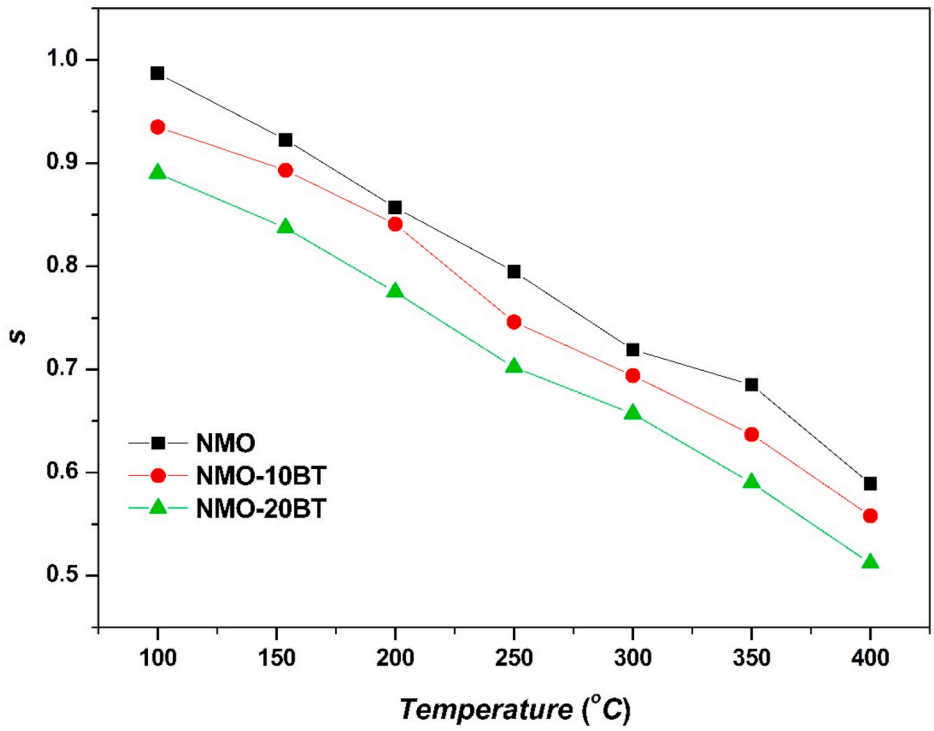


Fig. 9. Variation of power exponent s with temperature for NMO and NMO-BT composites.

maximum magnetization reduces on the addition of BT in NMO and NMO-20BT composites exhibits lowest maximum magnetization. The addition of non-magnetic BT creates spatial modifications in NMO and effects the movements of domains walls, subsequently lowering the maximum magnetization in the composites [40]. The remnant magnetization is also observed to decline with BT content in NMO-BT composites. However, in case of NMO-20BT composite, remnancy is more than that of NMO-10BT composite due to structural stability. The inset of Fig. 10 shows the enlarged image of *M-H* lines at lower magnetic fields. This improved magnetization shows potential for memory device applications. The coercivity of the NMO and NMO-BT composite are same which indicate that with addition of BT, the intrinsic magnetic behavior of NMO remains unaffected.

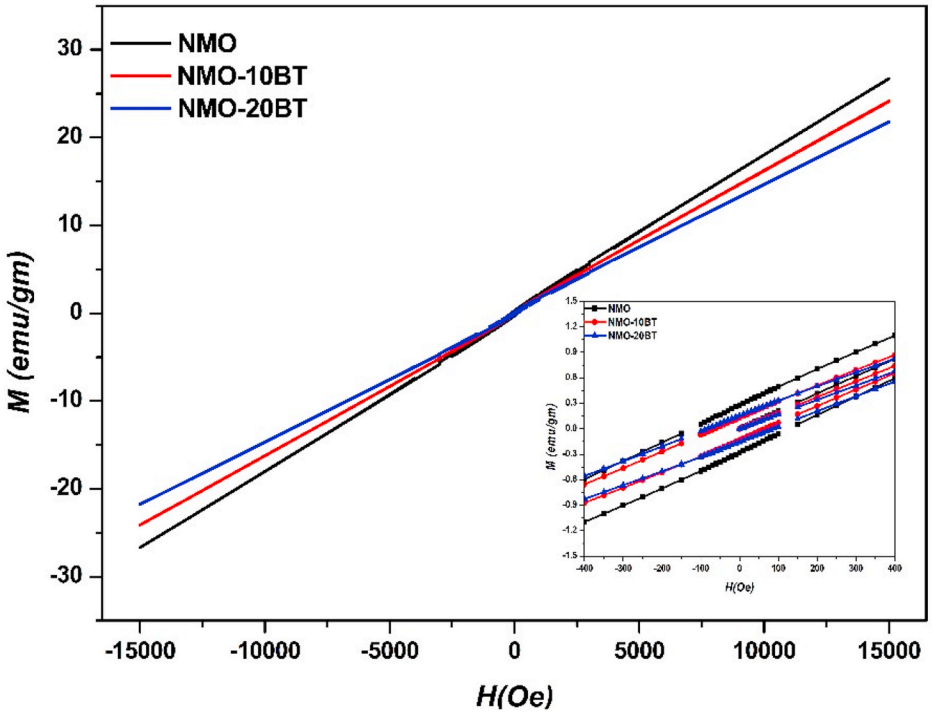


Fig. 10. M − H hysteresis loops of NMO and NMO-BT composites. Inset shows enlarged M − H hysteresis loops of NMO and NMO-BT composites in the magnetic field range of ±400 Oe.

3.6. Ferroelectric and magnetoelectric measurements

The ferroelectric properties of dielectrics are useful for devices based on polarization switching. The ferroelectric *P-E* hysteresis loop is studied using conventional Sawyer–Tower circuit [41] and the respective variations for NMO and NMO-20BT composite are plotted in Fig. 11. The *P-E* loop of NMO exhibits a low-lying flat loop suggesting its weak ferroelectric nature. It is interesting to note that the *P-E* loop of NMO-20BT composite display an improved saturation as well as polarization. This improvement can be attributed to infusion of ferroelectric domain walls in an orthorhombic lattice of NMO results in decrease in the lattice strain, which stabilizes the lattice. This decrement in strain stabilizes as well as saturates the ferroelectric hysteresis loop. NMO displayed an energy storage density of 2.004 mJ/cm^3 with efficiency of 17.21 % [32]. The NMO-20BT composite displays an enhanced energy storage density of 1.54 mJ/cm^3 with an enhanced efficiency of 50 %. This enhancement in energy storage efficiency of NMO-20BT composite confirms increase in the charge carriers movement and validates its potential for the electrical energy storage devices.

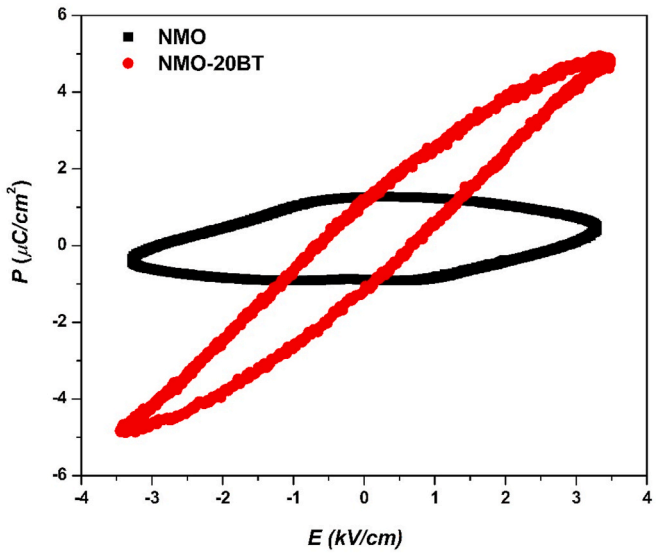


Fig. 11. P–E hysteresis loops of NMO and NMO-20BT composite.

The presence of magnetoelectric coupling in NMO-BT composites can be detected by studying the variation of ferroelectric *P-E* hysteresis loop with external dc magnetic field (*H*) ranging from 0 to 5000 Oe. The ferroelectric loops for NMO and NMO-BT composites were observed to vary with the applied magnetic field (not shown here). The measured maximum polarization (P_{max}) at respective magnetic fields for NMO, NMO-BT composites is shown in Fig. 12. The relative variation of 93 % in P_{max} versus *H* is maximum for NMO, which decreases up to 46 % for NMO-20BT composite. The blue line represents the fitted spectra for variation of NMO-BT composites. As the amount of BT is increased the percentage change of magnetoelectric coupling decreases indicating that the stabilization in the lattice decreases the strain originated from coupling amongst individual component phases. The magnetoelectric coupling coefficient was computed by non-linear fitting of variation of maximum polarization (P_{max}) with magnetic field (*H*). The error bars obtained from respective fitting are also incorporated in the experimental data plot. The maximum magnetoelectric coupling of 22 mV/cm/Oe was achieved for NMO-20BT composite.

4. Conclusions

Solid state ceramic method was employed for synthesizing $NdMnO_3$-$BaTiO_3$ (NMO-BT) composites. The influence of incorporation of BT in NMO on the structural, dielectric, impedance, magnetoelectric and multiferroic properties have been studied. The structural analysis from XRD confirm the composite formation in NMO-BT composites without any impurity phase. The NMO-BT composite sample exhibited giant dielectric constant with enhanced impedance, saturated ferroelectric behavior and modest remnant magnetization amongst the prepared composite samples. This improvement in electrical behavior of

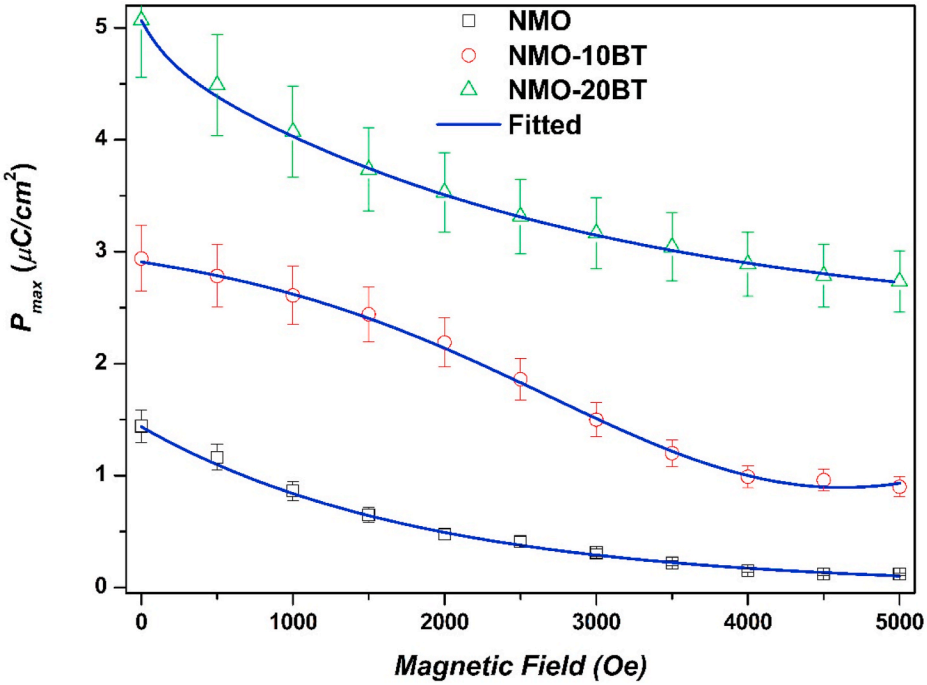


Fig. 12. Variation of maximum polarization versus magnetic field for NMO and NMO-BT composites.

$0.8NdMNO_3$-$0.2BaTiO_3$ (NMO-20BT) composite is due to lattice stability, thermally activated hopping conduction mechanism and highly resistive grains originating on account of BT addition. These results affirm that these NMO-BT composites exhibiting strong electrical, ferroelectric, magnetic and magnetoelectric characteristics and promise use of NMO-20BT composite for multiferroic magnetoelectric sensing devices.

4.1. Submission declaration

The authors (S·S, M·J and O·P·T) declare that this work described has been submitted and awarded as part of Ph.D thesis of Mr. S. Shankar (S·S) at Delhi Technological University, Delhi, India under the supervision of Dr. M. Jayasimhadri (M·J) and Dr. O. P. Thakur (O·P·T) and has not been published previously, that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder.

CRediT authorship contribution statement

S. Shankar: Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing, Visualization, Prof. **O.P. Thakur:** Writing – review & editing, Visualization, Supervision. **M. Jayasimhadri:** Writing – review & editing, Visualization, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is funded by Research Scheme no. 03(1427)/18/EMR-II, CSIR, New Delhi, India and supported by USIC, University of Delhi, New Delhi, India for characterization facilities.

References

[1] M. Kumar, S. Shankar, Brijmohan, S. Kumar, O.P. Thakur, A.K. Ghosh, Impedance spectroscopy and conductivity analysis of multiferroic BFO–BT solid solutions, Phys. Lett. 381 (2017) 379–386, https://doi.org/10.1016/j.physleta.2016.11.009.
[2] M.K. Kim, J.Y. Moon, S.H. Oh, D.G. Oh, Y.J. Choi, N. Lee, Strong magnetoelectric coupling in mixed ferrimagnetic-multiferroic phases of a double perovskite, Sci. Rep. 9 (2019) 1–10, https://doi.org/10.1038/s41598-019-41990-9.
[3] P. Zhou, A. V Singh, Z. Li, M.A. Popov, Y. Liu, D.A. Filippov, T. Zhang, W. Zhang, P. J. Shah, B.M. Howe, M.E. Mcconney, G. Srinivasan, M.R. Page, A. Gupta, Magnetoelectric interactions in composites of ferrite films on lattice-matched substrates and ferroelectrics, Phys. Rev. Appl. (2019), 054045, https://doi.org/10.1103/PhysRevApplied.11.054045, 1–15.
[4] W. Eerenstein, N.D. Mathur, J.F. Scott, Multiferroic and magnetoelectric materials, Nature 442 (2006) 759–765, https://doi.org/10.1038/nature05023.
[5] N.A. Spaldin, Multiferroics: past, present, and future, MRS Bull. 42 (2017) 385–390, https://doi.org/10.1557/mrs.2017.86.
[6] S. Lather, A. Gupta, J. Dalal, V. Verma, R. Tripathi, A. Ohlan, Effect of mechanical milling on structural, dielectric and magnetic properties of BaTiO3–Ni0.5Co0.5Fe2O4 multiferroic nanocomposites, Ceram. Int. 43 (2017) 3246–3251, https://doi.org/10.1016/j.ceramint.2016.11.152.
[7] M.D. Chermahini, M.M. Shahraki, M. Kazazi, Multiferroic properties of novel lead-free KNN-LT/20NZCFO magneto-electric composites, Mater. Lett. 233 (2018) 188–190, https://doi.org/10.1016/j.matlet.2018.09.001.
[8] D. Zhang, J. Cheng, J. Chai, J. Deng, R. Ren, Y. Su, H. Wang, C. Ma, C.S. Lee, W. Zhang, G.P. Zheng, M. Cao, Magnetic-field-induced dielectric behaviors and magneto-electrical coupling of multiferroic compounds containing cobalt ferrite/barium calcium titanate composite fibers, J. Alloys Compd. 740 (2018) 1067–1076, https://doi.org/10.1016/j.jallcom.2018.01.081.
[9] S. Liu, S. Yan, H. Luo, S. Huang, C. Liao, L. Deng, Magnetic effects on polarization response in particulate magnetoelectric Bi0.5Na0.5TiO3-La0.67Sr0.33MnO3 composites, Mater. Lett. 212 (2018) 139–142, https://doi.org/10.1016/j.matlet.2017.10.086.
[10] P.V.K.G.D. Jadav, C.S.L. Govindaraj, S.A.J.A. Bhalodia, Detailed investigations on structural properties and transport mechanism governed in - Nd 1 − x Cd x MnO 3, J. Mater. Sci. Mater. Electron. 29 (2018) 8107–8134, https://doi.org/10.1007/s10854-018-8817-6.
[11] B. Vyas, H. Kundalia, M. Udeshi, P. Trivedi, M. Vagadia, S. Rayaprol, Structure and magnetic behavior of Zn doped NdMnO 3 manganite : neutron diffraction study, Ceram. Int. 43 (2017) 14962–14967, https://doi.org/10.1016/j.ceramint.2017.08.015.
[12] G. Lalitha, P.V. Reddy, Elastic behavior of neodymium based manganites, Ultrasonics 52 (2012) 706–711, https://doi.org/10.1016/j.ultras.2012.01.006.
[13] S. Chanda, S. Saha, A. Dutta, T.P. Sinha, Raman spectroscopy and dielectric properties of nanoceramic NdFeO 3, Mater. Res. Bull. 48 (2013) 1688–1693, https://doi.org/10.1016/j.materresbull.2012.12.075.

[14] L. Kovács, K. Lengyel, M. Gospodinov, Antiferromagnetic ordering in RMnO3 rare-earth manganites observed by optical spectroscopy, Mater. Res. Bull. 46 (2011) 2567–2569, https://doi.org/10.1016/j.materresbull.2011.08.019.
[15] J.M.D. Coey, M. Viret, S. von Molnár, Mixed-valence manganites, Adv. Phys. 48 (1999) 167–293, https://doi.org/10.1080/000187399243455.
[16] T. Kimura, S. Ishihara, H. Shintani, T. Arima, T. Takahashi, K. Ishizaka, Y. Tokura, Distorted perovskite with eg1 configuration as a frustrated spin system, Phys. Rev. B Condens. Matter 68 (2003), 060403, https://doi.org/10.1103/PhysRevB.68.060403, 1–4.
[17] M. Udeshi, S. V Katba, A. Ravalia, M. Vagadia, Studies on the dielectric behavior of Cu-doped NdMnO3, AIP Conf. Proc. 1591 (2014) 1303–1305, https://doi.org/10.1063/1.4872938.
[18] A. Somvanshi, S. Husain, Study of structural , dielectric and optical properties of NdMnO3, AIP Conf. Proc. 3 (2018), 030242, https://doi.org/10.1063/1.5032577, 1–4.
[19] J.S.S.N. Aliouane, O. Prokhnenko, R. Feyerherm, M. Mostovoy, D.N.A.K. Habicht, K.C. Rule, E. Dudzik, A.U.B. Wolter, A. Maljuk, Magnetic order and ferroelectricity in RMnO3 multiferroic manganites: coupling between R- and Mn-spins, J. Phys. Condens. Matter 20 (2008) 434215.
[20] B. Bouadjemi, S. Bentata, A. Abbad, W. Benstaali, Ab-initio study of optoelectronic and magnetic properties of the orthorhombic NdMnO 3 perovskite, Solid State Commun. 207 (2015) 9–15, https://doi.org/10.1016/j.ssc.2015.02.001.
[21] W. Hu, G. Zhou, W. Ren, D. Li, Z. Zhang, Magnetic properties and unusual exchange coupling in self-organized NdMnO3/Mn3O4 nanocomposite films Magnetic properties and unusual exchange coupling in self-organized, J. Appl. Phys. 110 (2012), 013904, https://doi.org/10.1063/1.3596810.
[22] M. Faisal, A.A. Ismail, Doping strontium into neodymium manganites nanocomposites for enhanced visible light driven photocatalysis, Sci. Rep. 9 (2019) 1–11, https://doi.org/10.1038/s41598-019-50393-9.
[23] T. V Anuradha, Template-assisted sol-gel synthesis of nanocrystalline BaTiO 3, E-Journal Chem. 7 (2010) 894–898, https://doi.org/10.1155/2010/456263.
[24] F. Li, M. Zeng, H. Yu, H. Xu, J. Li, Dielectric characteristics of B-site-modified hexagonal-barium titanate, J. Mater. Sci. Mater. Electron. 27 (2016) 2836–2840, https://doi.org/10.1007/s10854-015-4098-5.
[25] M. Atif, M. Nadeem, W. Khalid, Z. Ali, Structural, magnetic and impedance spectroscopy analysis of (0.7)CoFe2O4+(0.3)BaTiO3 magnetoelectric composite, Mater. Res. Bull. 107 (2018) 171–179, https://doi.org/10.1016/j.materresbull.2018.07.026.
[26] B.D. Cullity, Elements of X-Ray Diffraction, second ed., Addison-Wesley Publishing Company Inc, New York, 1978.
[27] A. Jain, A.K. Panwar, A.K. Jha, Significant enhancement in structural, dielectric, piezoelectric and ferromagnetic properties of Ba0.9Sr0.1Zr0.1Ti0.9O3-CoFe2O4 multiferroic composites, Mater. Res. Bull. 100 (2018) 367–376, https://doi.org/10.1016/j.materresbull.2017.12.054.
[28] M. Kumar, S. Shankar, O.P. Thakur, A.K. Ghosh, Effects of co-substitution on dielectric, magnetic properties and magnetoelectric coupling in nano CoFe2O4, Mater. Lett. 143 (2015) 241–243, https://doi.org/10.1016/j.matlet.2014.12.101.
[29] S. Švarcová, K. Wiik, J. Tolchard, H.J.M. Bouwmeester, T. Grande, Structural instability of cubic perovskite BaxSr1 - xCo1 - yFeyO3 - δ, Solid State Ionics 178 (2008) 1787–1791, https://doi.org/10.1016/j.ssi.2007.11.031.
[30] S. Saha, S. Chanda, A. Dutta, T.P. Sinha, Dielectric relaxation of NdMnO3 nanoparticles, Mater. Res. Bull. 48 (2013) 4917–4923, https://doi.org/10.1016/j.materresbull.2013.07.027.
[31] S. Joseph, R.E. John, K.V. Saban, Unraveling the effects Mn-site doping on structural, magnetic and magnetotransport properties of Nd0.67Sr0.33Mn0.9T0.1O3 (T= Mn, Fe, Cr, Ni), Ceram. Int. 45 (2019) 21249–21262, https://doi.org/10.1016/j.ceramint.2019.07.107.
[32] S. Shankar, M. Kumar, V. Tuli, O.P. Thakur, M. Jayasimhadri, Energy storage and magnetoelectric coupling in ferroelectric–ferrite composites, J. Mater. Sci. Mater. Electron. 29 (2018) 18352–18357, https://doi.org/10.1007/s10854-018-9949-4.
[33] T. Chakraborty, R. Yadav, S. Elizabeth, H.L. Bhat, Evolution of Jahn-Teller distortion, transport and dielectric properties with doping in perovskite NdFe1-xMnxO3 ($0 \leq x \leq 1$) compounds, Phys. Chem. Chem. Phys. 18 (2016) 5316–5323, https://doi.org/10.1039/c5cp07083j.
[34] M. George, S.S. Nair, K.A. Malini, P.A. Joy, M.R. Anantharaman, Finite size effects on the electrical properties of sol-gel synthesized CoFe2O4 powders: deviation from Maxwell-Wagner theory and evidence of surface polarization effects, J. Phys. D Appl. Phys. 40 (2007) 1593–1602, https://doi.org/10.1088/0022-3727/40/6/001.
[35] A. Kaur, L. Singh, K. Asokan, Electrical relaxation and conduction mechanisms in iron doped barium strontium titanate, Ceram. Int. 44 (2018) 3751–3759, https://doi.org/10.1016/j.ceramint.2017.11.158.
[36] M. Kumar, S. Shankar, R.K. Kotnala, Om Parkash, Evidences of magneto-electric coupling in BFO-BT solid solutions, J. Alloys Compd. 577 (2013) 222–227, https://doi.org/10.1016/j.jallcom.2013.04.044.
[37] S. Shankar, O.P. Thakur, M. Jayasimhadri, Conductivity behavior and impedance studies in BaTiO3–CoFe2O4 magnetoelectric composites, Mater. Chem. Phys. 234 (2019) 110–121, https://doi.org/10.1016/j.matchemphys.2019.05.095.
[38] S. Shankar, O.P. Thakur, M. Jayasimhadri, Impedance spectroscopy and conduction behavior in CoFe2O4-BaTiO3 composites, J. Electron. Mater. (2019), https://doi.org/10.1007/s11664-019-07700-x.
[39] S. Shankar, M. Kumar, A.K. Ghosh, O.P. Thakur, Conduction mechanism and dielectric properties of BiFeO3–BaTiO3 solid solutions, J. Mater. Sci. Mater. Electron. 25 (2014) 4896–4901, https://doi.org/10.1007/s10854-014-2250-2.
[40] S. Shankar, M. Kumar, A.K. Ghosh, O.P. Thakur, M. Jayasimhadri, Anomalous ferroelectricity and strong magnetoelectric coupling in CoFe2O4-ferroelectric composites, J. Alloys Compd. 779 (2019) 918–925, https://doi.org/10.1016/j.jallcom.2018.11.252.
[41] A. Peláiz-Barranco, J.D.S. Guerra, O. García-Zaldívar, F. Calderón-Piñar, M. E. Mendoza, D.A. Hall, E.B. Araújo, Phase transition and dielectric properties of La-doped Pb(Zr, Ti)O3 antiferroelectric ceramics, Solid State Commun. 149 (2009) 1308–1311, https://doi.org/10.1016/j.ssc.2009.05.004.

Materials Today: Proceedings xxx (xxxx) xxx

Contents lists available at ScienceDirect

Materials Today: Proceedings

journal homepage: www.elsevier.com/locate/matpr

Structural and photoluminescence properties of Er^{3+} doped $SrBi_2Nb_2O_9$ ceramics

Ritushree Shaily, Renuka Bokolia *

Department of Applied Physics, Delhi Technological University, Delhi, India

ARTICLE INFO

Article history:
Received 31 March 2021
Received in revised form 23 May 2021
Accepted 24 May 2021
Available online xxxx

Keywords:
Photoluminescence
Ferroelectrics
Solid state method
X-Ray diffraction
Orthorhombic distortion
Ceramics

ABSTRACT

In the present study, structural and photoluminescence properties of polycrystalline Er^{3+} doped $SrBi_{2-x}Nb_2Er_xO_9$ (x = 0.00, 0.01, 0.02, 0.03, 0.04, 0.05) ceramics were investigated. All the samples were prepared by the conventional solid-state reaction method and were sintered at 1000 ℃ for 8 h. X-Ray Diffraction (XRD) confirms the formation of single-phase material with orthorhombic structure. An increase in the lattice parameters and unit cell volume is observed with increasing Erbium (Er^{3+}) content. Strong green emission at 550 nm is seen under 480 nm excitation wavelength at room temperature for an optimum Er content (x = 0.03) and is attributed to $^4S_{3/2} \rightarrow {}^4I_{15/2}$ transitions.
© 2021 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the International Conference on Futuristic Research in Engineering Smart Materials.

1. Introduction

Photoluminescence (PL), an optical phenomenon is characterized by the emission of two or more photons of the longer wavelength low energy in the visible region means when a molecule absorbs a photon, so one of its electrons gets excited and jumps to the higher electronic excited state and so further the excited electron returns to its lower energy state followed by the emission of photons of lower energy. In rare-earth materials, since different ions have different energy level structure, based on that excitation light sources such as Ultraviolet, Visible, or Infra-red light is chosen. Recently, photoluminescence observed in rare-earth-doped materials has attracted a lot of attention due to its wide range of energy-related applications such as light and displays, temperature sensing, biomedical and solar cells etc. [1,2].

Ferroelectric materials, a type of materials that hold spontaneous polarization and are reversible when subjected to an electric field. Recently, ferroelectric materials have drawn so much attention because of their excellent properties such as high dielectric constant, great fatigue resistance, high Curie temperature, and amazing piezoelectric effects making them attractive materials for different applications such as non-volatile random-access memories (NvRAM), pyroelectric infrared detectors, high capacitance capacitors and optical switches etc. [3,4]. The bismuth layered structure ferroelectric (BLSFs) ceramics, member of Aurivillius family with a chemical formula: $(Bi_2O_2)^{2+}$ $(A_{n-1}B_nO_{3n+1})^{2-}$, where A can be a mono-, di-, or trivalent element (or their combination) with coordination no. 12, for instance Ba^{2+}, Ca^{2+}, La^{3+}, Sm^{3+}, Bi^{3+}, Sr^{2+}, etc., furthermore B is a transition element suited to octahedral coordination possessing coordination no. 6 such as Fe^{3+}, Nb^{5+}, Ti^{3+}, Ta^{5+}, W^{6+},and Mo^{6+}, etc. [5–8], and also, n denotes the number of perovskite layers intermediate between two bismuth oxide layers. Hereabouts, A could be Ca^{2+}, Pb^{2+}, Ba^{2+} and Sr^{2+}, etc. and B perhaps Ti^{4+}, Nb^{5+}, Ta^{5+}, V^{5+}, Mo^{5+}, and W^{6+} etc. Ferroelectric ceramics are very sensitive to compositional modifications such as by changing stoichiometry or even by doping ceramics with numerous elements notably modifying the ferroelectric properties and crystal structure of the ceramics.

$SrBi_2Nb_2O_9$ (SBN), lead-free ferroelectric member of Bismuth Layered Structure Ferroelectric Ceramics belonging to the Aurivillius family (n = 2) has drawn consideration due to their remarkable characteristics such as a high Curie temperature, T_c (718 K), great fatigue resistance, low dielectric loss, reasonable spontaneous polarization, and temperature-insensitive high piezoelectric coefficient and considered for high-temperature applications. Some rare-earth ions like Er^{3+}, Eu^{3+}, Pr^{3+}, etc. [9]. Are widely used as luminescence centres for their potential applications. Thus, by rare-earth doping, the electrical properties and strong luminescence of some ferroelectric materials gets enhanced. In this paper,

* Corresponding author.
E-mail address: renukabokolia@dtu.ac.in (R. Bokolia).

https://doi.org/10.1016/j.matpr.2021.05.520
2214-7853/© 2021 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the International Conference on Futuristic Research in Engineering Smart Materials.

Please cite this article as: R. Shaily and R. Bokolia, Structural and photoluminescence properties of Er^{3+} doped $SrBi_2Nb_2O_9$ ceramics, Materials Today: Proceedings, https://doi.org/10.1016/j.matpr.2021.05.520

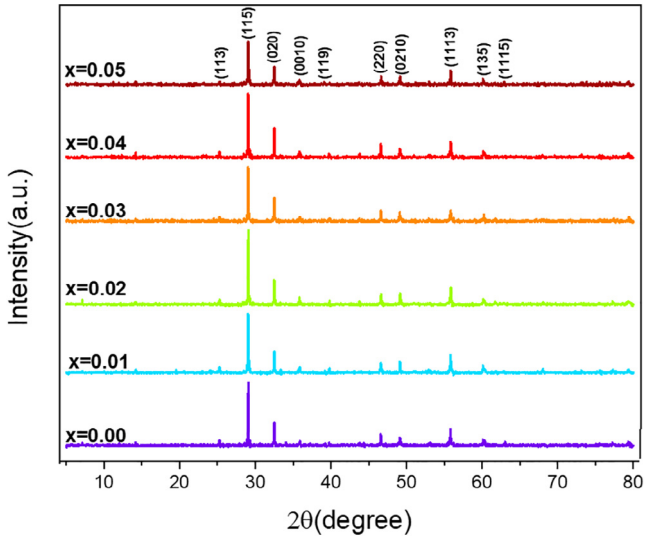


Fig. 1. The XRD patterns of pure and Er^{3+} doped $SrBi_2Nb_2O_9$ at different concentrations of Er^{3+} based on the formula $SrBi_{2-x}Nb_2Er_xO_9$ ceramics.

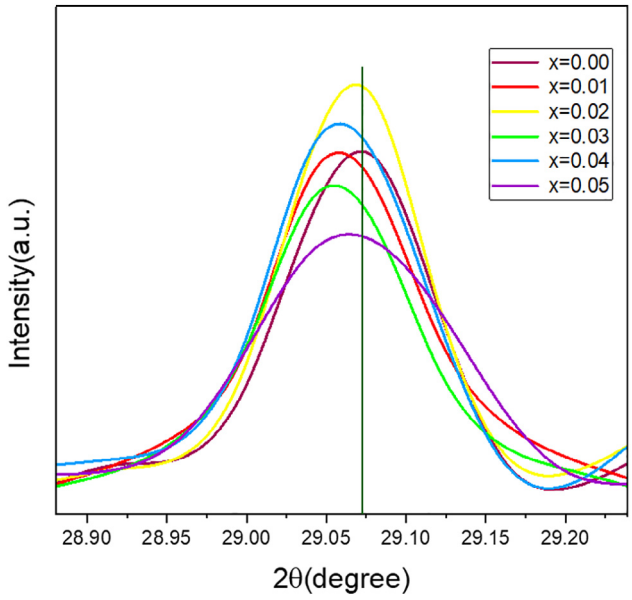


Fig. 2. Shift of strongest XRD peaks (1 1 5) at different Erbium concentrations.

we report the synthesis of Er^{3+} doped $SrBi_2Nb_2O_9$(SBN) ceramic by conventional solid-state method and study the associated effect of doping of Er^{3+} on their respective crystal structures and luminescence properties.

2. Material synthesis and characterization details

In this experiment, the polycrystalline Er^{3+} doped $SrBi_{2-x}Nb_2Er_xO_9$ (x = 0.00, 0.01, 0.02, 0.03, 0.04, 0.05) ceramics in stoichiometric compositions were synthesized using conventional solid-state reaction method. Firstly, highly purified starting materials Nb_2O_5 (from Aldrich of 99.9% purity), $SrCO_3$ (from Alfa Aesar of 97.5% purity), Bi_2O_3 (from Alfa Aesar of 97.5% purity), and Er_2O_3 (from Alfa Aesar of 97.5% purity) weighed in stoichiometric ratios and then these weighed samples were mixed. After that mixed sample was grinded into a mortar pestle with the addition of ethanol till it gets dried up and repeats the process for 6 h. Then the prepared samples were calcined at 950 ℃ for 7 h. After calcination, 5 wt% of polyvinyl alcohol (PVA), standard binder is added to the calcined samples and mixed rigorously and then disk-shaped green pellets of all samples were prepared with a diameter of 13 mm and thickness of 1 mm by manual hydraulic press at a pressure of 50 MPa, and keep at rest for 1 min. Hereafter, green pellets of all samples were sintered at 1000 ℃ for 8 h with intermediate heat treatment at 500 ℃ for 1 h in order to remove the binder from the pellets [10].

The phase identification and analysis of crystal structure of Er^{3+} doped $SrBi_{2-x}Nb_2Er_xO_9$ were done with the help of the X-Ray diffraction. The XRD spectra achieved by using the Bruker D-8 Advance X-ray diffractometer at Cu-Kα radiation (k_α = 1.5406 Å). The structural morphology and compositional analysis of pure $SrBi_2Nb_2O_9$ and Er^{3+} doped $SrBi_{2-x}Nb_2Er_xO_9$ have been investigated on unpolished sintered pellets by using the JEOL 661 Scanning electron microscopy (SEM) equipped with Amkette energy dispersive spectrometer (EDS). The photoluminescence (PL) spectra of Er^{3+} doped SBN recorded at an excitation wavelength, λ_{ex} = 480 nm.

3. Results and discussion

3.1. Structural analysis

Erbium, rare earth metal belongs to the lanthanide series shows most stable + 3 oxidation state thus it is easy to substitute Er^{+3} in a place of Bi^{+3} [11]. Fig. 1, shows XRD spectra of pure $SrBi_2Nb_2O_9$ and Er^{3+} doped $SrBi_{2-x}Nb_2Er_xO_9$ ceramic powders at varying Er^{3+} concentrations (x = 0.00, 0.01, 0.02, 0.03, 0.04, 0.05). From the XRD analysis in the range 2ϑ = 10–80, the observed Bragg's reflection was indexed and found closely matching with the standard diffraction pattern data of $SrBi2Nb_2O_9$ with JCPDS card#01–089- 8154 and confirmed the formation of single phase with no traces of secondary phases or unreacted phases present, specifying that Er^{3+} ions completely dissolved in SBN host lattice without any significant structural change of the host lattice. Thus, all samples crystallized in the orthorhombic phase, associated with the space group $A2_1am$. Also, the highest diffraction peak associated with (1 1 5) plane of $SrBi_2Nb_2O_9$ was detected for all concentrations conforming to the bismuth layered structure with n = 2 and the results are in agreement with the earlier reports indicating the strongest diffraction corresponding to (112n + 1) reflection in the Aurivillius phase for BLSF compositions. The lattice parameters were calcu-

Table 1
Lattice parameters of $SrBi_{2-x}Nb_2Er_xO_9$.

$SrBi_{2-x}Nb_2Er_xO_9$	a(Å)	b(Å)	c(Å)	Volume (Å^3)	Orthorhombic Distortion, b/a	Orthorhombicity, 2(a-b)/(a + b)	Tetragonal strain, c/a
x = 0.00	5.4933	5.5136	25.035	758.2565	1.003695	0.00369	4.55737
x = 0.01	5.4938	5.5141	25.039	758.5155	1.003695	0.00369	4.557683
x = 0.02	5.4943	5.5146	25.043	758.7745	1.003695	0.00369	4.557996
x = 0.03	5.4946	5.5152	25.047	759.0197	1.003749	0.00374	4.558476
x = 0.04	5.4949	5.5157	25.051	759.2512	1.003785	0.00378	4.558955
x = 0.05	5.4951	5.5162	25.055	759.4689	1.003840	0.00383	4.559517

Table 2
EDS compositional analysis.

Nominal composition	Sr (at%)	Bi (at%)	Nb (at%)	Er (at%)	Er/Bi
Stoichiometric $SrBi_2Nb_2O_9$	5.731	11.797	12.691	–	–
Present Work $SrBi_{2-x}Nb_2Er_xO_9$					
x					
0.00	3.53	11.151	6.363	–	–
0.01	2.652	8.328	4.775	0.034	0.004
0.02	2.653	8.319	4.777	0.068	0.008
0.03	2.654	8.253	4.779	0.103	0.012
0.04	2.656	8.215	4.782	0.137	0.016
0.05	2.657	8.177	4.784	0.172	0.021

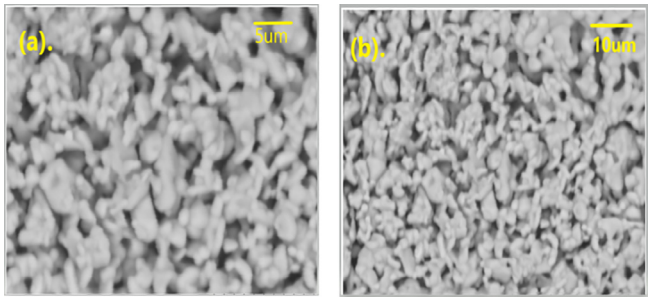


Fig. 3. Surface morphology of SBN ceramic.

lated with the help of PowderX software, and the structural parameters tabulated in Table 1.

It is noted from the Fig. 2. with an increase in the concentration of Er3 + ions the XRD peaks at 29.08 shifted towards lower 2θ values shows that the Er3 + ions completely diffused into the host lattice along with the increase in lattice parameters which is due to the difference in ionic radii of Bi3+(r = 102A) and Er3+ (r = 0.88A) leading to structural distortion and is well supported with the increasing elemental ratio of (Er/Bi) as seen through EDS compositional analysis (Table 2). Also, with the increase in the concentration of erbium ions there is an expansion in a volume of unit cell, an increase in the orthorhombic distortion and orthorhombicity followed by the increase in tetragonal strain. [3,9].

The microstructure of the sintered ceramic surface of $SrBi_{2-x}Nb_2Er_xO_9$ is shown in Fig. 3. The morphological details reveal the formation of randomly oriented grains of varying size with a presence of porosity in the microstructure as solid-state method leads to more pores and non-uniform grain size, and was achieved under prepared condition. The micrographs indicated insignificant changes in grain size with varying increasing Er content. The compositional analysis of pure $SrBi_2Nb_2O_9$ and Er3 + doped $SrBi_{2-x}Nb_2Er_xO_9$ pellets was done by EDS (Energy Dispersive Spectrum) and the data of the elements were given in the Table 2.

3.2. Photoluminescence spectra

The photoluminescence (PL) properties of Er^{3+} doped $SrBi_{2-x}Nb_2Er_xO_9$ (x = 0.00, 0.01, 0.02, 0.03, 0.04 and 0.05) ceramics for different concentrations were examined under the excitation wavelength, λ_{ex} = 480 nm at room temperature and acquired photoluminescence emission spectrum is shown in Fig. 4. From the observations, all the Er doped $SrBi_{2-x}Nb_2Er_xO_9$ ceramics display strong green emission ranging from 529.2 nm to 549.8 nm, (centred at 529.2 nm, 540.96 nm, 542.95 nm, and 549.8 nm) refers $^2H_{11/2}\rightarrow{}^4I_{15/2}$, and $^4S_{3/2}\rightarrow{}^4I_{15/2}$ transitions, respectively and light greenish emission was observed at 579.4 nm. Furthermore, the splitting of PL spectra ranging from 540.96 nm to 549.8 nm into

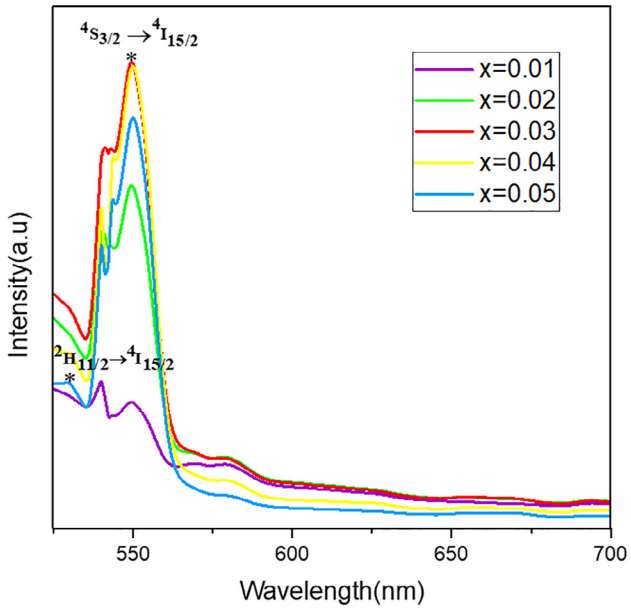


Fig. 4. PL spectra of Er^{3+} doped $SrBi_{2-x}Nb_2Er_xO_9$ at different concentrations under 480 nm excitation wavelength.

several components is due to Stark effect (or Stark splitting) [12,13]. The highest peak of PL emission spectra monitored at a wavelength of 549.8 nm i.e., the strong green emission refers to the excitation of Er3 + ions from ground level $^4I_{15/2}$ to higher energy level $^4S_{3/2}$ rather than the host material. With an increase in the doping content of Er3 + firstly, the PL intensity increases up to × = 0.03 and then decreases because by means of concentration quenching while the position of emission peaks for all remains the same. As indicated in PL spectra, the maximum emission intensity occurs at a concentration, x = 0.03. So, Er3 + doped $SrBi_{2-x}Nb_2Er_xO_9$ (x = 0.00, 0.01, 0.02, 0.03, 0.04 and 0.05) ceramics can be considered as appropriate visible light excited green phosphors where Er3 + doped $SrBi_{2-x}Nb_2Er_xO_9$ with concentration, x = 0.03 ceramics outshines others [14–16].

4. Conclusion

The synthesis of Er^{3+} doped $SrBi_2Nb_2O_9$ ceramics, their structural analysis and PL properties were investigated and studied. The study of XRD patterns of pure SBN and Er^{3+} doped SBN indicates that the pure SBN is orthorhombic in nature with no secondary phases and with further increase in the doping concentration of Er^{3+} ion in SBN results in the shifting of strongest peak (115) towards lower angles and there is an increase in orthorhombic distortion and orthorhombicity with the increase in Er content. The PL properties implies, at a 480 nm excitation

wavelength, strong green emission bands observed at 549.8 nm and strongest emission intensity was achieved at a concentration of Er^{3+} when x = 0.03.

CRediT authorship contribution statement

Ritushree Shaily: Conceptualization, Data curation, Investigation, Methodology, Writing - original draft. **Renuka Bokolia:** Conceptualization, Methodology, Supervision, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Authors would like to express gratitude towards Research Project Grant (F.NO. DTU/IRD/619/2019/2112), Delhi Technological University for granting research facilities. Authors wish to acknowledge the help provided by the technical and support staff in the Department of Applied Physics, Delhi Technological University for aiding with the characterization of samples.

References

[1] Y. Zhao, X. Wang, Y. Zhang, Y. Li, X. Yao, Optical temperature sensing of up-conversion luminescent materials: Fundamentals and progress, J. Alloy. Compd. 817 (2020), https://doi.org/10.1016/j.jallcom.2019.152691 152691.

[2] D.P. Volanti, I.L.V. Rosa, E.C. Paris, C.A. Paskocimas, P.S. Pizani, J.A. Varela, E. Longo, The role of the Eu3+ ions in structure and photoluminescence properties of SrBi2Nb2O9 powders, Opt. Mater. (Amst) 31 (6) (2009) 995–999, https://doi.org/10.1016/j.optmat.2008.11.006.

[3] T. Wei et al., Enhanced up-conversion photoluminescence and dielectric properties of Er- and Zr-codoped strontium bismuth niobate ceramics, Ceram. Int. 41 (9) (2015) 12364–12370, https://doi.org/10.1016/j.ceramint.2015.06.067.

[4] D. Peng, H. Sun, X. Wang, J. Zhang, M. Tang, X. Yao, Red emission in Pr doped CaBi4Ti4O15 ferroelectric ceramics, Mater. Sci. Eng. B Solid-State Mater. Adv. Technol. 176 (18) (2011) 1513–1516, https://doi.org/10.1016/j.mseb.2011.09.009.

[5] V. Senthil, T. Badapanda, A. Chandra Bose, S. Panigrahi, Enhancement of dielectric and ferroelectric properties of dysprosium substituted SrBi2Ta2O9 ceramics, J. Mater. Sci. Mater. Electron. 27 (2) (2016) 1602–1608, https://doi.org/10.1007/s10854-015-3930-2.

[6] S. Jain, P. Ganguly, S. Devi, A.K. Jha, Structural, dielectric and ferroelectric studies of molybdenum substituted Sr2Bi2Nb2O9 ferroelectric ceramics, Ferroelectrics 381 (1) (2009) 152–159, https://doi.org/10.1080/00150190902870051.

[7] O. Jalled, Z. Alharbi, S.R. Alharbi, A. Saeed, M. Alhassan, S. Al-Heniti, H.Y. Mohammed, Y. Al-Hadeethi, F. Al-Marzouki, A. Al-Mujtaba, Synthesis and dielectric properties of nanocrystalline strontium bismuth niobate, J. Nanosci. Nanotechnol. 17 (1) (2017) 594–600, https://doi.org/10.1166/jnn.2017.12459.

[8] L. Yu, J. Hao, Z. Xu, W. Li, R. Chu, G. Li, Strong photoluminescence and good electrical properties in Eu-modified SrBi2Nb2O9 multifunctional ceramics, Ceram. Int. 42 (13) (2016) 14849–14854, https://doi.org/10.1016/j.ceramint.2016.06.119.

[9] T. Wei, C.Z. Zhao, Q.J. Zhou, Z.P. Li, Y.Q. Wang, L.S. Zhang, Bright green upconversion emission and enhanced ferroelectric polarization in Sr1-1.5xErxBi2Nb 2O9, Opt. Mater. (Amst) 36 (7) (2014) 1209–1212, https://doi.org/10.1016/j.optmat.2014.03.001.

[10] A. Banwal, R. Bokolia, Phase evolution and microstructure of BaBi2Nb2O9 ferroelectric ceramics, Mater. Today Proc. vol. 3, no. xxxx (2020) 2–5, https://doi.org/10.1016/j.matpr.2020.09.380.

[11] P. Henderson, "Rare earth element geochemistry, Developments in geochemistry," vol. 2, p. 510, 1984, [Online]. Available: http://ezproxy.lib.utexas.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=geh&AN=1984-023694&site=ehost-live.

[12] A. Tomar, M. Singh, S. Singh, L. Sharma, S. Arya, S. Kasana, Ultraviolet Quantum Cutting through down Conversion Luminescence Behaviour of Er3+ Substituted Sr0.7Bi2.2Nb2O9 (BLFS) Ceramics, Integr. Ferroelectr. 204 (1) (2020) 33–37, https://doi.org/10.1080/10584587.2019.1674983.

[13] T. Honma, T. Komatsu, D. Zhao, H. Jain, Writing of rare-earth ion doped lithium niobate line patterns in glass by laser scanning, IOP Conf. Ser. Mater. Sci. Eng. 1 (2009), https://doi.org/10.1088/1757-8981/1/1/012006 012006.

[14] L. Mukhopadhyay, V.K. Rai, R. Bokolia, K. Sreenivas, 980 nm excited Er3+/Yb3+/Li+/Ba2+: NaZnPO4 upconverting phosphors in optical thermometry, J. Lumin. 187 (October) (2017) 368–377, https://doi.org/10.1016/j.jlumin.2017.03.035.

[15] R. Bokolia, O.P. Thakur, V.K. Rai, S.K. Sharma, K. Sreenivas, Dielectric, ferroelectric and photoluminescence properties of Er3+ doped Bi4Ti3O12 ferroelectric ceramics, Ceram. Int. 41 (4) (2015) 6055–6066, https://doi.org/10.1016/j.ceramint.2015.01.062.

[16] R. Bokolia, V.K. Rai, L. Chauhan, K. Sreenivas, Structural and light up-conversion luminescence properties of Er3+-Yb3+-W6+ substituted Bi4Ti3O12, AIP Conf. Proc. 1731 (2016) 4–7, https://doi.org/10.1063/1.4948208.

Materials Today: Proceedings xxx (xxxx) xxx

Contents lists available at ScienceDirect

Materials Today: Proceedings

journal homepage: www.elsevier.com/locate/matpr

Structural properties of Strontium Bismuth Niobate ($SrBi_2Nb_2O_9$) ferroelectric ceramics

Pooja Mojumdar, Ritushree Shaily, Renuka Bokolia *

Department of Applied Physics, Delhi Technological University, New Delhi, India

ARTICLE INFO

Article history:
Received 31 March 2021
Received in revised form 22 May 2021
Accepted 24 May 2021
Available online xxxx

Keywords:
SBN
Phase evolution
XRD
Ferroelectric
Raman
Ceramics

ABSTRACT

The pure $SrBi_2Nb_2O_9$ (SBN) ceramic was synthesized by the conventional solid-state reaction method. The sample was then undergoing heat treatment at different calcination temperatures: 800 °C, 850 °C, and 950 °C, and the evolution of phases were studied at these different calcination temperatures, followed by sintering at 1000 °C of temperature to form single-phase SBN. Few secondary phases have been identified along with the desired SBN phase but begin to decrease with the increase in the calcination temperature. The most appropriate calcination and sintering temperature have been successfully investigated with no impurity phase. The structural morphology of the sample sintered at 1000 °C studied under SEM reveals the formation of the highly dense plate-like grains and EDS provides information about the elemental composition. The Raman analysis revealed the formation of the orthorhombic phase of pure SBN.
© 2021 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the International Conference on Futuristic Research in Engineering Smart Materials.

1. Introduction

Ferroelectric ceramics were the backbone of billion-dollar industries from the high-dielectric capacitor to positive temperature coefficient devices and electro-optic valves because they have great features such as high dielectric permittivity, outstanding piezoelectric response, and swappable macroscopic polarization [1,2]. The goal to achieve better electronic device technology and higher efficient energy production devices was the main focus for rapid development in ferroelectric oxide ceramics. Oxides with Perovskite structure have general formula ABO_3 where A is a large positive ion with 12-fold coordinate and B is a small negative ion with 6-fold coordinated with oxygen anions have been broadly studied in the search of new material for device application [3]. After a detailed study of ferroelectric, we found BSLF (Bismuth layered ferroelectric) has gained much momentum for its properties such as high curie temperature coefficient, low ageing rate, and electro-optic behaviour. BSLF has great potential in electronic devices. BSLF are from the Aurivillius family having formula $(Bi_2O_2)^{2+}(A_{n-1}B_nO_{3n+1})^{2-}$ where $(Bi_2O_2)^{2+}$ is bismuth layer and $(A_{n-1}B_nO_{3n+1})^{2-}$ is pseudo perovskite layer. The A site is supposed to be occupied by cation Na^+, K^+, Sr^{2+}, Ca^{2+}, Ba^{2+}, Ba^{3+} and B site is occupied by W^{6+}, Nb^{5+}, Ti^4, etc. [4]. Bismuth layer ferroelectric having n = 2 i.e., $CaBi_2Nb_2O_9$ (CBN), $BaBi_2Nb_2O_9$ (BBN), $SrBi_2Nb_2O_9$(SBN) have many properties like the low operating voltage, excellent switching speed, small leakage current density. Two-layer Aurivillius compound (n = 2) has a high curie temperature $\geq$ 900 °C so they have the best application in spacecraft, aerospace, and nuclear plants [5]. The two-layer Aurivillius family has one of the most important ceramic compounds Strontium Bismuth Niobate ($SrBi_2Nb_2O_9$). $SrBi_2Nb_2O_9$ is a BSLF material that has special importance due to its anti- fatigue property, spontaneous polarization, and low-temperature coefficient of resonance frequency [6]. $SrBi_2Nb_2O_9$ has great application in non-volatile random-access memory and in fine frequency tolerance resonators [7–8]. SBN is a lead-free relaxor ceramic. All these properties have inspired us to synthesize Strontium Bismuth Niobate. In the present, we study variations in phase evolution at different calcination temperatures during the preparation of $SrBi_2Nb_2O_9$ ceramic to get a single-phase SBN with no impurity phase.

2. Experimental details

2.1. Synthesis

Pure Strontium Bismuth Niobate ($SrBi_2Nb_2O_9$) sample was synthesized by the conventional solid-state method [9]. Firstly, high

* Corresponding author.
E-mail address: renukabokolia@dtu.ac.in (R. Bokolia).

https://doi.org/10.1016/j.matpr.2021.05.522
2214-7853/© 2021 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the International Conference on Futuristic Research in Engineering Smart Materials.

Please cite this article as: P. Mojumdar, R. Shaily and R. Bokolia, Structural properties of Strontium Bismuth Niobate ($SrBi_2Nb_2O_9$) ferroelectric ceramics, Materials Today: Proceedings, https://doi.org/10.1016/j.matpr.2021.05.522

/ce:cross-ref>. Firstly, high purity raw materials: $SrCO_3$ (>99.9%), Nb_2O_5 (>99.9%), and Bi_2O_3 (>99.7%) were stoichiometrically weighed and then mixed. The mixture of powders was grinded for 6 h followed by calcination. The grinded powder was calcinated at three different calcination temperatures: 800 °C, 850 °C, and 950 °C with a hold of 3 h. The calcined powder was rigorously mixed with 5% wt. polyvinyl alcohol, a binder, and then the resulting powder was converted into disc-shaped pellets of thickness 1 mm with a diameter of 13 mm manually by the hydraulic press at a pressure of 50 MPa with 1 min resting time. After that, the pellets were finally sintered at 1000 °C for 8 h with intermediate heat treatment at 500 ℃ for 1 h so that there will be no traces of binder left in the sample. The equation for synthesis:

$$SrCO_3 + Bi_2O_3 + Nb_2O_5 \rightarrow SrBi_2Nb_2O_9 + CO_2 \uparrow$$

2.2. Characterization techniques

The crystallographic structural analysis of pure SBN was studied under X-Ray diffraction (XRD, Bruker D8 Advance X-ray diffractometer) using Cu-Kα radiation (λ = 0.154 nm). The XRD spectrum was observed within the range, $20° \leq 2\vartheta \leq 70°$. The peak of ceramics was matched with JCPDF file number 01–089-8154 to confirm the pure phase of SBN. The crystalline size has been calculated by Scherrer's equation:

$$D = \frac{0.89\lambda}{\beta cos\theta}$$

where, D is the average crystalline size, λ is Cu-Kα_1(λ = 0.154 nm) wavelength, β is the full width at half maximum (FWHM) of the diffraction peak, and θ is the angle of diffraction.

The structural morphology and grain formation have been investigated on unpolished sintered ceramic pellets by using the JEOL 661 Scanning electron microscopy (SEM) equipped with Amkette energy dispersive spectrometer (EDS). At room temperature, the Raman spectra were recorded by using the Horiba LabRam HR evolution spectrometer at wavelength 514 nm at 40 mW.

3. Result and discussion

3.1. X-Ray diffraction analysis

Fig. 1(a, b, c) depicts the evolution of phase that develops during the calcinations heat treatment at 800 °C, 850 °C, and 950 °C of $SrBi_2Nb_2O_9$(SBN) powder to form a single-phase of SBN. The calci-

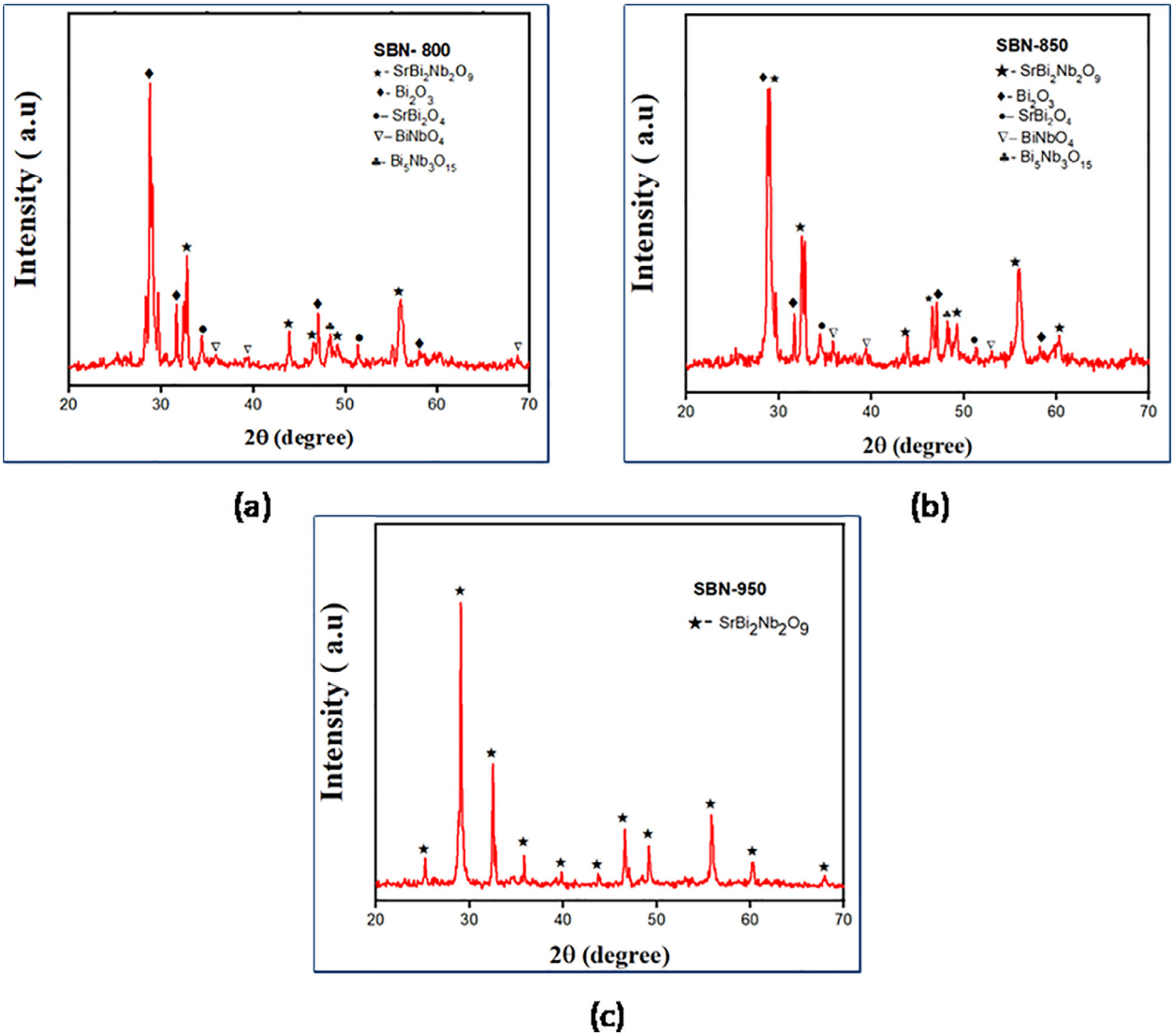


Fig. 1. XRD pattern of $SrBi_2Nb_2O_9$ at different temperature (a) at 800 °C (b) 850 °C (c) 950 °C.

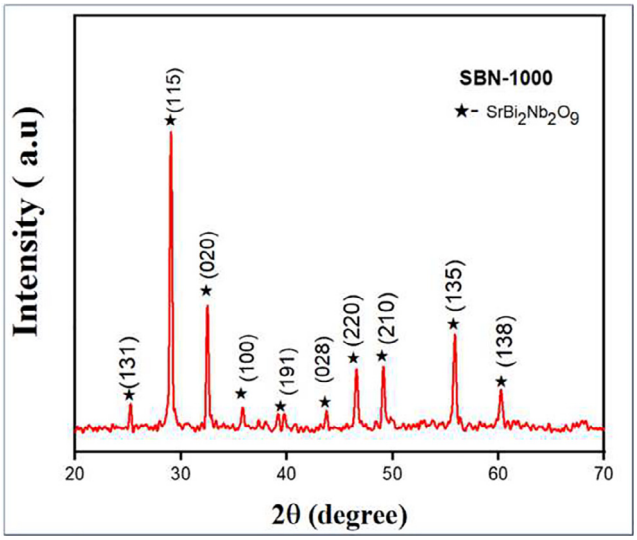


Fig. 2. XRD pattern of $SrBi_2Nb_2O_9$ ceramic pellet sintered at temperature 1000 °C.

nation temperature uprises to 1000 °C to ensure that single-phase SBN ($SrBi_2Nb_2O_9$) is formed. Highly pure raw materials $SrCO_3$, Nb_2O_5, and Bi_2O_3 were mixed stoichiometrically, as the reaction between them proceeds to lead to the formation of various intermediate phases accompanying the initial crystallization of $SrBi_2Nb_2O_9$ phase. During the reaction between Bi_2O_3 and Nb_2O_5 numerous phases e.g., $Bi_8Nb_{18}O_{57}$, $BiNbO_4$ and $Bi_5Nb_3O_{15}$ are formed [10]. At 800 °C calcined temperature, the XRD pattern shows the presence of Bi_2O_3, $SrBi_2O_4$ (due to the reaction between Bi_2O_3 and $SrCO_3$), single-phase $Bi_5Nb_3O_{15}$, and some intermediate phase of $BiNbO_4$, existed side by side with desired single-phase $SrBi_2Nb_2O_9$ phase and the same intermediate phases were also detected at 850 ℃. At 850 °C, the XRD pattern shows the mix phase of $SrBi_2Nb_2O_9$ and Bi_2O_3 at 28.82° & 29.06° respectively. From the literature survey, we found that peaks corresponding to Bi_2O_3 phase seen at 2θ (31.66°, 47.08°, and 58.06°) because Bi_2O_3 remains unreacted but when it is treated to a higher range of temperature, the phase of Bi_2O_3 disappears. Also, it is interesting to see that intermediate phases i.e., $SrBi_2O_4$, $Bi_5Nb_3O_{15}$, and $BiNbO_4$ start decreasing with the increase in the calcination temperature. When the calcination temperature rises to 950 ℃ no secondary phases are observed for the sample indicating that the reaction mechanism for the phase evolution of pure SBN is completed followed by sin-

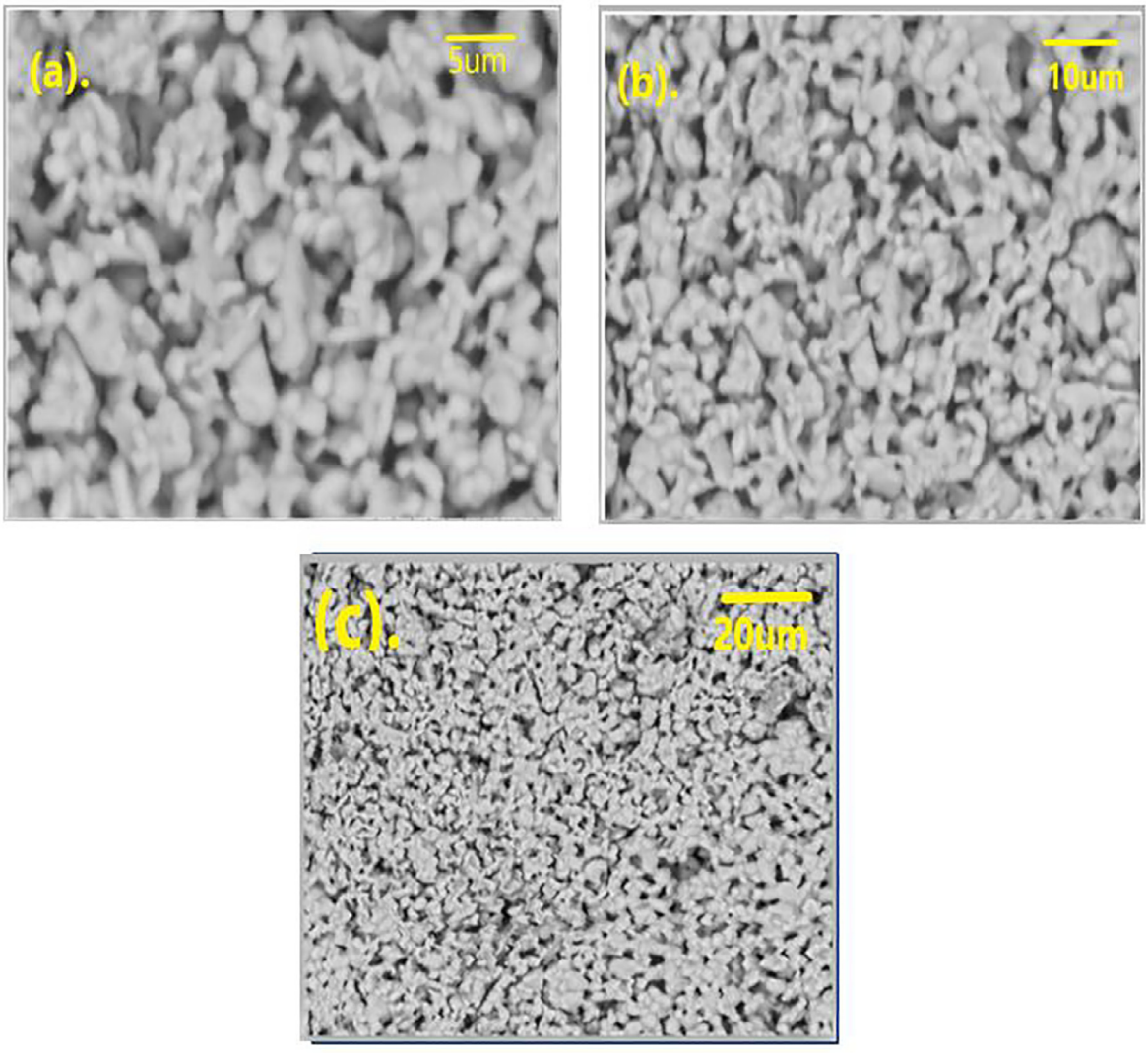


Fig. 3. SEM image of pure SBN.

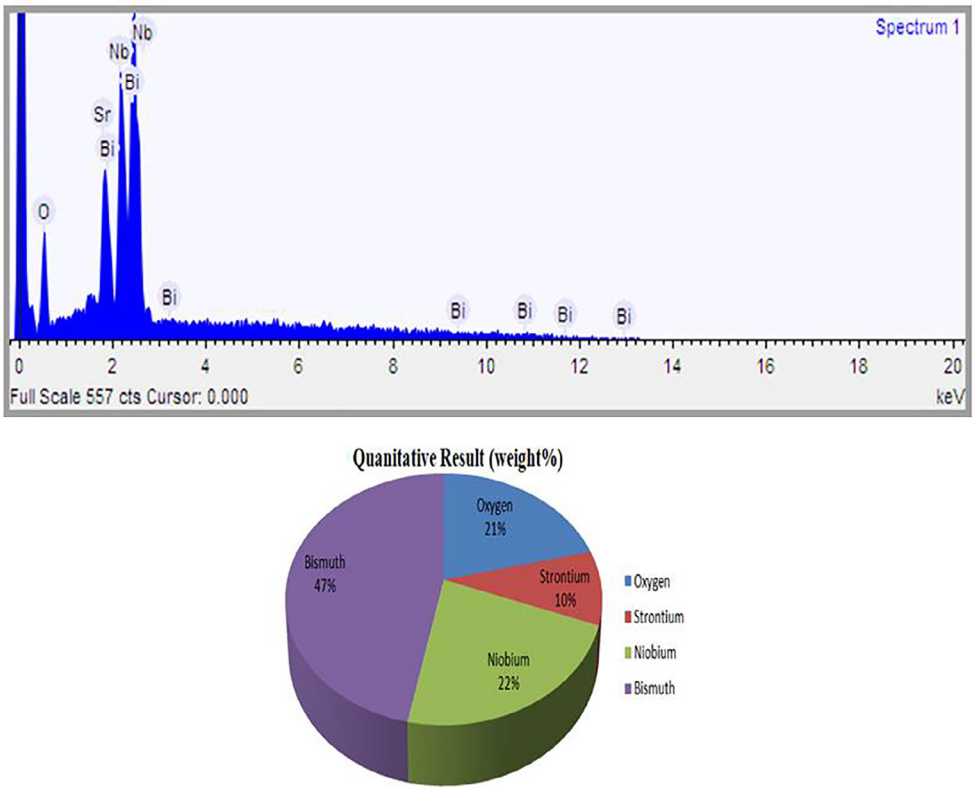


Fig. 4. EDS graph of pure SBN.

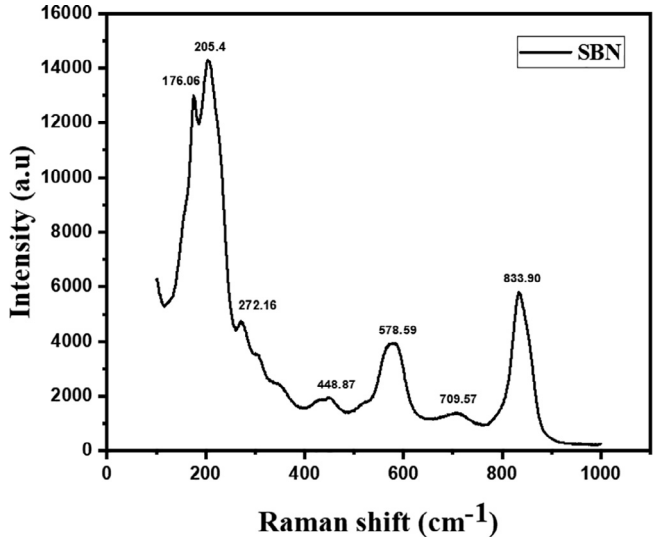


Fig. 5. Raman Spectra of SBN ceramics at spectral range 100 cm^{-1} to 1000 cm^{-1}.

tering at 1000 °C. The peak of ceramics is matched with JCPDF file number 01-089-8154 confirms the pure SBN of single-phase and no secondary phases were detected. Thus, the reaction mechanism for obtaining the desire phase of $SrBi_2Nb_2O_9$, should be calcinated at 950 °C and sintered at 1000 °C. (Fig. 2). The lattice parameters for pure SBN a = 5.4933 Å, b = 5.5136 Å, and c = 25.035 Å related to the orthorhombic phase, associated with the space group $A2_1am$. The lattice parameter was calculated with the help of PowderX software. Also, the highest diffraction peak associated with the (1 1 5) plane of $SrBi_2Nb_2O_9$ was detected. The average crystallite size of SBN was calculated by Scherrer's equation which is found to be 31.50 nm. As the intensity and sharpness of the peak increases with an increase in temperature attributed to increase in crystallinity.

Table 1
Composition of elements present in pure SBN, sintered at 1000 °C.

Element	Weight %	Weight % (σ)	Atomic %
Oxygen	21.212	0.983	69.781
Strontium	9.541	0.544	5.731
Niobium	22.402	0.868	12.691
Bismuth	46.844	1.154	11.797

3.2. SEM and EDS analysis

Morphology and compositional analysis of synthesized pure SBN have been investigated along with the Energy-dispersive spectrum (EDS). Fig. 3 shows SEM image of pure SBN ceramic pellet with their EDS graph. The micrograph shows the formation of grains that are randomly oriented with a presence of porosity in the microstructure as the solid-state method leads to more pores and non-uniform grain size, and the dense sample was achieved underprepared conditions. Porosity leads to conductivity and has some application such as thermal stability, high permeability, and high resistance to chemical attack. Moreover, due to high-density growth, agglomerations in the nanoparticles were also seen [11,12]. The energy dispersive spectroscopy (EDS) was used for elemental analysis and from the analysis the information obtained was that no other impurity elements were detected confirms the absence of secondary phases in the required sintered sample and those detected elements: O, Sr, Bi, and Nb were present in the desired composition in the sample, described in Table 1. (See Fig. 4.)

3.3. Raman scattering study

Raman Spectroscopy, chemical analysis technique based on the interaction of light with the chemical bonds present within a material. Raman spectroscopy is utilized for investigating the basic lattice symmetry, crystal structure, and molecular interactions of ferroelectric compounds. Fig. 5 shows the Raman Spectra of sintered SrBi2Nb2O9 (pure SBN) ceramic pellet, at room temperature within the spectral range 100 cm^{-1} to 1000 cm^{-1}. Some intense peaks can be seen at 205.4 cm^{-1}, 578.59 cm^{-1} and 833.90 cm^{-1} with relatively large Raman intensity, apart from that some weak peaks are also observed at 176.06 cm^{-1}, 272.16 cm^{-1}, 448.87 cm^{-1} and 709.57 cm^{-1}. The 176.06 cm^{-1} mode corresponds to Nb z-axis vibrations, modes at 205.4 cm^{-1} attributed to the Bi^{3+} vibrations or related to the vibrations of A-site ions. The mode at 272.16 cm^{-1} is leading by a force constant due to Bi-O_2 and Nb-Bi-O_2 bonds. The mode at 449.87 cm^{-1} moved to low frequency for scanty Sr. The mode at 709.57 cm^{-1} assigns to asymmetric deformation of BO_3 triangular group and distortion of BO_4 tetrahedron Also, mode at 578.59 cm^{-1} corresponds to a rigid sub-lattice mode where the displacement of positive and negative ion displacement is equal and opposite. The mode located at 833.90 cm^{-1} reflects symmetric stretching of the octahedral NbO_6 and the presence of NbO_6 clearly indicates the predominance of the orthorhombic phase. So, at 176.06 cm^{-1} 205.4 cm^{-1}, 578.59 cm^{-1}, and 833.90 cm^{-1} modes related to the orthorhombic phase which is in agreement with the XRD pattern [13–16].

4. Conclusion

Pure SBN ceramic prepared by the conventional solid-state method was investigated to study the effect of different calcination temperatures: 800 °C, 850℃, and 950 °C, on the evolution of phases during the single-phase formation of pure SBN. XRD spectra reveal with an increase in the calcination temperature, the intensity of intermediate phases decreases and at 950℃ no impurity phases were detected and show the formation of peaks of a single phase of $SrBi_2Nb_2O_9$ with orthorhombic geometry, followed by the sintering of the sample at 1000 °C. The Surface micrograph shows the formation of grains that are randomly oriented with a presence of porosity in the microstructure and the elemental analysis by EDS reveals that only elements O, Sr, Bi, and Nb were detected in the desired composition. The Raman study shows the SBN orthorhombic phase and the modes located at 176.06 cm^{-1} 205.4 cm^{-1}, 578.59 cm^{-1}, and 833.90 cm^{-1} related to the orthorhombic phase of pure SBN.

CRediT authorship contribution statement

Pooja Mojumdar: Conceptualization, Methodology, Investigation, Data curation, Writing - original draft. **Ritushree Shaily:** Conceptualization, Methodology, Investigation, Data curation, Writing - original draft. **Renuka Bokolia:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing, Supervision.

Declaration of Competing Interest

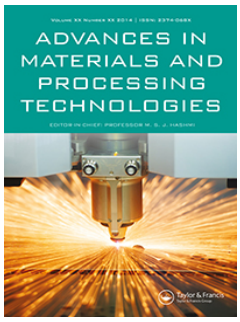
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

I would like to thank Research Project Grant (F.NO. DTU/IRD/619/2019/2112) Delhi Technological University for providing the research facilities and acknowledged technical staff, AIC, Department of Applied Physics for aiding with the characterization of the samples.

References

[1] G.H. Haertling, <1999 Ferroelectrics, history and technology, Haertling-1999-Journal_of_the_American_Ceramic_Society.pdf>, J. Am. Ceram. Soc. 82 (1999) 797–818.
[2] J.F. Scott, Applications of modern ferroelectrics, Science (80-.) 315 (5814) (2007) 954–959, https://doi.org/10.1126/science.1129564.
[3] M.A. Peña, J.L.G. Fierro, Chemical structures and performance of perovskite oxides, Chem. Rev. 101 (7) (2001) 1981–2017, https://doi.org/10.1021/cr980129f.
[4] Z. Peng, D. Yan, Q. Chen, D. Xin, D. Liu, D. Xiao, J. Zhu, Crystal structure, dielectric and piezoelectric properties of Ta/W codoped Bi3TiNbO9 Aurivillius phase ceramics, Curr. Appl. Phys. 14 (12) (2014) 1861–1866, https://doi.org/10.1016/j.cap.2014.10.011.
[5] S. Zhang, F. Yu, Piezoelectric materials for high temperature sensors, J. Am. Ceram. Soc. 94 (10) (2011) 3153–3170, https://doi.org/10.1111/j.1551-2916.2011.04792x.
[6] P.H. ABELSON, Materials research, Science (80-.) 235 (4784) (1987) 9, https://doi.org/10.1126/science:235.4784.9.
[7] P. Yang, D.L. Carroll, J. Ballato, R.W. Schwartz, Growth and optical properties of SrBi2Nb2O9 ferroelectric thin films using pulsed laser deposition, J. Appl. Phys. 93 (11) (2003) 9226–9230, https://doi.org/10.1063/1.1571219.
[8] A. Ando, T. Sawada, H. Ogawa, M. Kimura, Y. Sakabe, Fine-tolerance resonator applications of bismuth-layer-structured ferroelectric ceramics, Japanese J. Appl. Physics, Part 1 Regul. Pap. Short Notes Rev. Pap. 41 (Part 1, No. 11B) (2002) 7057–7061, https://doi.org/10.1143/JJAP.41.7057.
[9] A. Banwal, R. Bokolia, Phase evolution and microstructure of BaBi2Nb2O9 ferroelectric ceramics, Mater. Today Proc. vol. 3, no. xxxx (2020) 2–5, https://doi.org/10.1016/j.matpr.2020.09.380.
[10] M. Verma, A. Tanwar, K. Sreenivas, Phase evolution of strontium bismuth niobate ceramics by conventional solid-state reaction method, J. Therm. Anal. Calorim. 135 (4) (2019) 2077–2087, https://doi.org/10.1007/s10973-018-7287-7.
[11] D.P. Volanti, I.L.V. Rosa, E.C. Paris, C.A. Paskocimas, P.S. Pizani, J.A. Varela, E. Longo, The role of the Eu3+ ions in structure and photoluminescence properties of SrBi2Nb2O9 powders, Opt. Mater. (Amst) 31 (6) (2009) 995–999, https://doi.org/10.1016/j.optmat.2008.11.006.
[12] M.K. Adak, A. Mukherjee, A. Chowdhury, J. Khatun, U.K. Ghorai, D. Dhak, Electrical and energy storage properties of nickel substituted barium bismuth niobate nano-ceramics prepared by chemical route, J. Mater. Sci. Mater. Electron. 29 (18) (2018) 15847–15858, https://doi.org/10.1007/s10854-018-9671-2.
[13] D.P. Volanti, L.S. Cavalcante, E.C. Paris, A.Z. Simões, D. Keyson, V.M. Longo, A.T. de Figueiredo, E. Longo, J.A. Varela, F.S. De Vicente, A.C. Hernandes, Photoluminescent behavior of Sr Bi2 Nb2 O9 powders explained by means of B- Bi2 O3 phase, Appl. Phys. Lett. 90 (26) (2007) 261913, https://doi.org/10.1063/1.2753114.
[14] G.Z. Liu, C. Wang, H.S. Gu, H. Bin Lu, Raman scattering study of La-doped SrBi2Nb2O 9 ceramics, J. Phys. D Appl. Phys. 40 (24) (2007) 7817–7820, https://doi.org/10.1088/0022-3727/40/24/034.
[15] R. Singh, V. Luthra, R.S. Rawat, R.P. Tandon, Structural, dielectric and piezoelectric properties of SrBi2Nb2O9 and Sr0.8Bi2.2Nb2O9 ceramics, Ceram. Int. 41 (3) (2015) 4468–4478, https://doi.org/10.1016/j.ceramint.2014.11.139.
[16] H.C. Gupta, Archana, V. Luthra, Lattice vibrations of ABi2Nb2O9 crystals (A = Ca, Sr, Ba), Vib. Spectrosc. 56 (2) (2011) 235–240, https://doi.org/10.1016/j.vibspec.2011.03.002.



Advances in Materials and Processing Technologies

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tmpt20

Study of residual stresses in multi-pass friction stir processed surface composites

Kartikeya Bector, Mrinal Singh, Divya Pandey, Ravi Butola & Ranganath M. Singari

To cite this article: Kartikeya Bector, Mrinal Singh, Divya Pandey, Ravi Butola & Ranganath M. Singari (2021): Study of residual stresses in multi-pass friction stir processed surface composites, Advances in Materials and Processing Technologies, DOI: 10.1080/2374068X.2021.1939983

To link to this article: https://doi.org/10.1080/2374068X.2021.1939983

Published online: 15 Jun 2021.

Submit your article to this journal

Article views: 26

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tmpt20

ADVANCES IN MATERIALS AND PROCESSING TECHNOLOGIES
https://doi.org/10.1080/2374068X.2021.1939983

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

Check for updates

Study of residual stresses in multi-pass friction stir processed surface composites

Kartikeya Bector, Mrinal Singh, Divya Pandey, Ravi Butola and Ranganath M. Singari

Department of Mechanical Engineering, Delhi Technological University, Delhi, India

ABSTRACT

Friction stir processing (FSP) has materialised as a significantly improved procedure for surface modification and fabrication. This has propelled its use in the industries of aerospace, automobile, and suchlike. Surface composites have many advantages over conventional composites and FSP is extensively used for their fabrication. In this study, we have compared and analysed the residual stresses of such friction stir fabricated surface composites with AA6063 as the base alloy and B_4C nanoparticles as the reinforcement. Here, diethyl ether was used for the incorporation of B_4C into the substrate of the base alloy. The residual stresses were measured using a portable stress analyser. The effect of the number of passes (one pass, two passes, three passes, and four passes) of FSP was studied. It was concluded that the specimen processed at 1000 rpm and three passes showed the best results.

ARTICLE HISTORY
Accepted 4 June 2021

KEYWORDS
friction stir processing; residual stress; surface composite; multi-pass; boron carbide

1. Introduction

Friction stir processing (FSP) is a technique established by Mishra et al. [1] in 2002 through which metal matrix composites (MMCs) can be developed and surface modification can be carried out. It is unlike other fabrication techniques as it does not require any melting of the material; rather, the whole process is done in the solidstate itself. In FSP, frictional heat generated due to the contact of the tool and the workpiece softens and plasticises the material and alters the microstructure of the surface of the metals. It can also be utilised to develop surface composites (SCs) with the incorporation of reinforcement onto the surface of metals. The resultant composite retains the bulk properties of the metal along with the surface properties of the reinforcement. Such surface composites have found widespread applications in aerospace, defence, and other such industries.

FSP has shown to considerably enhance the mechanical, microstructural, and tribological properties of metal matrices. FSP is a derivative process of FS Welding (FSW). Rajendran et al. [2] compared the load-carrying capabilities of FSWed aluminium joints in AA2014-T6 alloy with those of TIG-welded and riveted ones. FSWed joints showed upto 75% higher load-carrying capacity than its counterparts. Kurtyka et al. [3] showed that FSP increased hardness and compressive strength in A339/SiC composite. In

CONTACT Ravi Butola ravibutola33855@gmail.com Department of Mechanical Engineering, Delhi Technological University, Delhi, 110042

© 2021 Informa UK Limited, trading as Taylor & Francis Group

AA6063, it led to the improvement of ductility and tensile strength [4]. Chaudhary et al. [5] summarised the efficacy of FSP for surface refinement and the resulting improvement of the mechanical and tribological properties of SCs of other alloys. Butola et al. [6,7] studied the use of FSP with natural reinforcements. Tyagi et al. [8] studied the hybrid composition of ceramic particles and aloevera ash as FSP reinforcements. Khayyamin et al. [9] investigated the microstructure of FSPed AZ91/SiO_2 composite and found a homogeneous dispersion of reinforcement and smaller particle size, thus improving the tensile strength. FSPed composites were also shown to be more wear resistant and had a smaller value of friction coefficient than the base metal [10]. The choice of reinforcement particles also affects the resulting properties in the surface composite. Hardness and wear tests were performed on different composites having different reinforcement materials, including SiC, TiC, ZrO_2, and B_4C [11]. Rajendran et al. [12] determined the optimum processing parameters of FSW to maximise the tensile strength of AA2014-T6 butt joints using RSM. They [13] also studied the impact of varying the tool tilt angle from 0° to 4° vis-à-vis the microstructural characteristics and strength of of the FSWed joint. They found the joint at 2° to be the most suitable one. Butola et al. [14,15] used B_4C nanoparticles with AA7075-T6 base alloy for the production of SCs by FSP via the Self-Assembled Monolayer technique. Other extrinsic factors such as tool rotation and traversal speeds, number of passes, tool tilt angle, tool pin profile, and others were also studied and their effects noted [16–20].

Residual stresses refer to those stresses which remain in the material even after an external load has been removed. Residual stress is undesirable and can significantly impact the integrity and performance of the composite. Mehdi and Mishra [21] used Pulstec XRD portable machine to study the TIG+FSPed joint of two distinct aluminium alloys. It has been shown that FSP decreases residual stresses in surface composites [22,23]. Generally, more residual stresses have been observed on the advancing side than the retreating side [24]. Woo et al. [25] detailed the dominant factor for residual stresses in surface composites as the frictional heat from the tool shoulder. The welding speed and tool rotational speed were also shown to affect the residual stress generated in the surface composites [26].

The use of volatile diethyl ether (DEE) for the bonding of B_4C nanoparticles into the grooves in the surface of the AA6063 substrate. In this paper we further study and analyse the influence of the number of passes on residual stress in the AA6063/B_4C surface composites.

2. Experimentation procedure

2.1. Materials and method

With the rising demands of industries such as aerospace and automotive, aluminium matrix composites (AMCs), having superior qualities such as high wear and corrosion resistance as well as good surface finish and high strength-weight ratio, are need of the hour [27]. AA6063 aluminium alloy is an Al–Mg–Si-based alloy of the 6xxx series family of aluminium alloys, with mechanical properties similar to the pure aluminium metal [28]. The Al-B_4C metal matrix composites (MMCs) have, particularly, gained popularity in in the nuclear industry for their ability in absorbing thermal neutrons [29].

In this study, the base metal (BM) comprises AA6063 plates (chemical composition shown in Table 1), while the B_4C nanoparticles (<30 nm) act as the reinforcement. In order to place these B_4C particles, a groove of length 160 mm, 2 mm width, and a 3.5 mm depth was carved in the BM plate (dimensions – 200 mm × 80 mm × 6 mm) for the placement of reinforcement. A viscous paste of B_4C and diethyl ether was formed, which was then applied to the grooves before friction stir processing of these grooves was done. H13 tool steel was used to make the tool for this experiment.

Keeping the number of passes as the variable process parameter, the samples were processed at 1000 rpm tool rotation speed and 40 mm/min tool traverse speed using a square pin profile (Figure 1). The samples were processed at one pass, two passes, three passes, and four passes (Table 2).

2.2. Fabrication of surface composite

During friction stir processing, a non-consumable, high speed, rotating tool with a pin and shoulder is plunged into our sample piece and then rotated in the desired direction. During this exchange, a certain amount of heat is produced in the interface of the tool shoulder and the sample piece, softening the material [30].

This, along with the mechanical stirring by the pin, leads to the intense plastic deformation of the material. Since FSP is a relatively new method, several studies have tried to find the relation between the different microstructural and tribological properties. Ghasemi-Kahrizsangi et al. [31] studied the FSP of plates of mild steel and inferred that the number of passes and rotation speed had a direct relationship with the improvement of mechanical properties. A simplified diagram of the FSP arrangement, as well as the actual apparatus involved, is shown side by side in Fig. 2(a) and Fig. 2(b), respectively.

2.3. Residual stress testing

In the modified material, after friction stir processing, certain forces known as residual stresses remain in the composite even after the withdrawal of the rotating tool. These stresses have a prominent influence on sample longevity, because of which measurement of such stresses becomes all the more important. However, traditional methods such as numerical simulation have not succeeded in providing with the correct data mainly due to the problems involved in analysis of such small spot diameters [32]. In recent times, this has given rise to development and popularity of a powerful destructive tool procedure called X-ray diffraction method (XRD). This method is based on Bragg's equation, that is a function of interplanar distance D and the diffraction angle.

Table 1. Chemical composition of AA6063 base alloy.

Element	Al	O	Si
Wt %	92.37	6.21	1.42

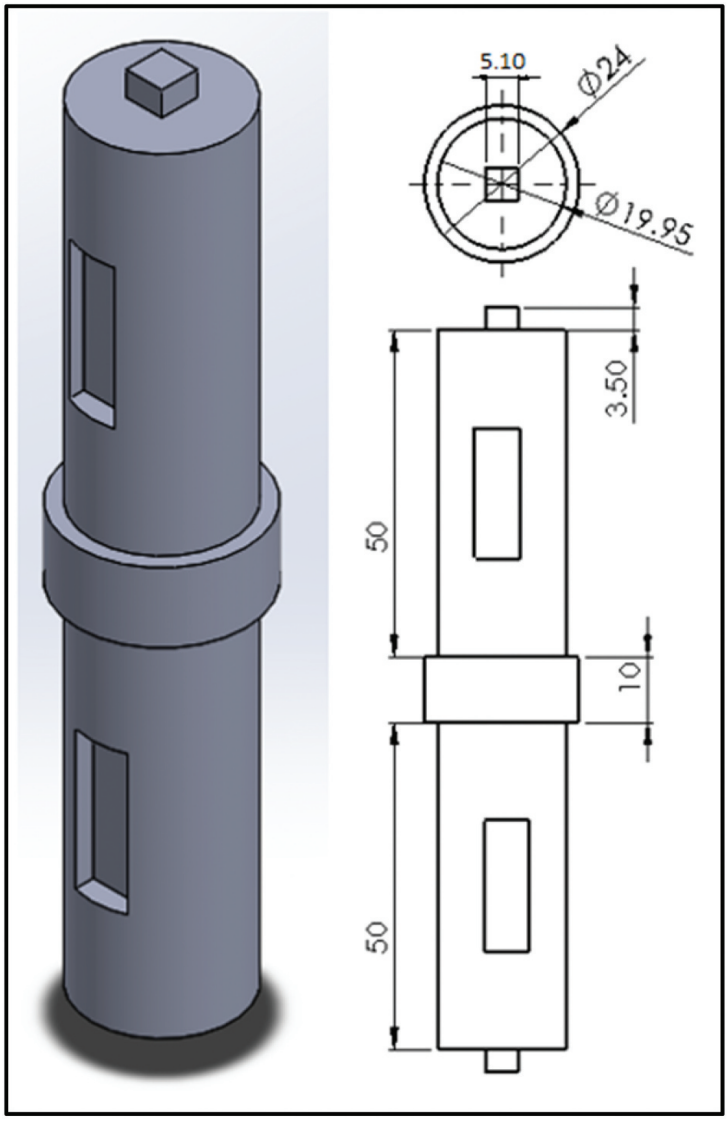


Figure 1. Square tool pin.

This method has further been subdivided into $\sin^2\psi$ and cos method. The former is a relatively older method in which interplanar strain and spacing in the plane (hkl) are measured, followed by the calculation of residual stress using the change in angle [33]. In this, the theory of elasticity states that the required stress can be calculated as

$$\varepsilon^{\varphi\psi} = \frac{1+\nu}{E}\left(\sigma_x cos^2\phi + \tau_{xy} sin2\phi + \sigma_y sin2\phi - \sigma_z\right) sin^2\psi + \frac{1+\nu}{E}\left(\tau_{xz} cos\phi + \tau_{yz} sin\phi\right) sin2\psi + \frac{1+\nu}{E}\sigma_z - \frac{\nu}{E}\left(\sigma_x + \sigma_y + \sigma_z\right) \quad (1)$$

Where

E – Elastic modulus

ν – Poisson's ratio

σ_1 – Maximum stress

σ_2 _ Minimum stress

Φ and ψ – Azimuth angles

Table 2. Process parameters.

Sample number	Number of passes	Tool rotational speed
S1	1	1000
S2	2	1000
S3	3	1000
S4	4	1000

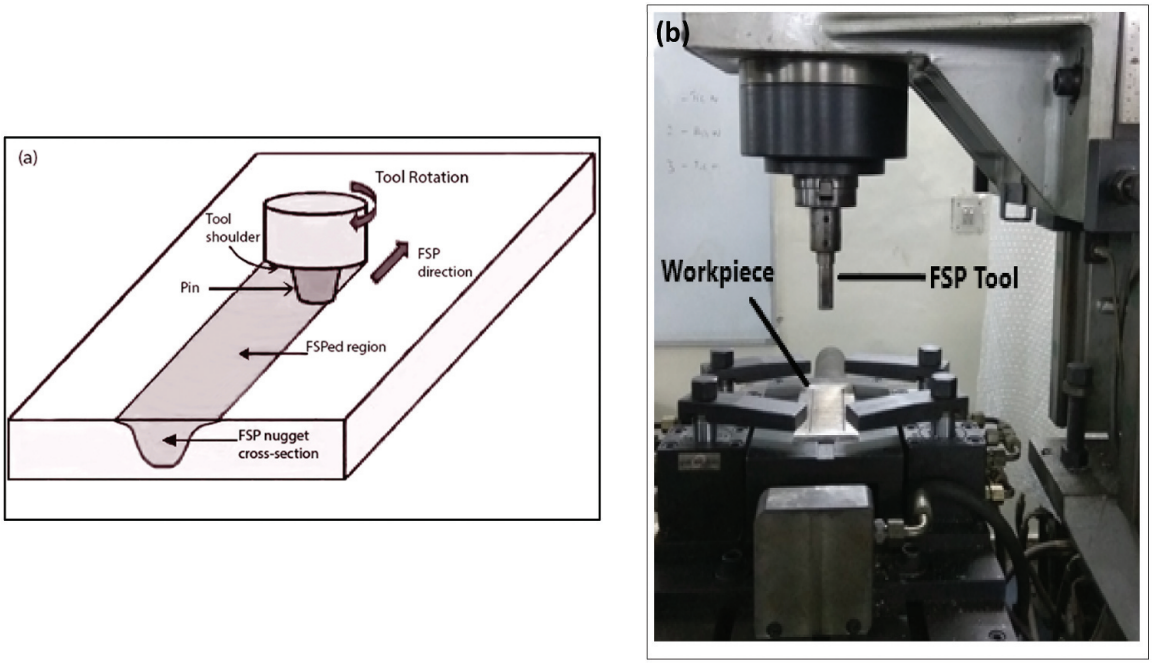


Figure 2. **(a)** Schematic diagram of FSP; **(b)** experimental setup of FSP.

ψ – Angle between incident and the normal line of the surface of the sample

Φ – Angle between the projection of incidence on the plane of the sample and the X axis

$\varepsilon^{\Phi\psi}$ – Elastic strain of materials along the incident direction [34].

However, due to weak X-ray penetration, one can only measure the residual stress using the above expression. Further, considering the stress on the surface to be 2D and the normal stress to be zero, the partial derivation of $\sin^2\psi$ is found to be

$$\sigma = K\frac{\partial^2\theta}{\partial^2 sin^2\psi} \tag{2}$$

$$K = -\frac{E}{2(1+\nu)}\cot\left(\theta\frac{\pi}{180}\right) \tag{3}$$

Where

K – X-ray stress constant

$\frac{\partial^2\theta}{\partial^2\sin^2\Psi}$ – Measured experimentally

The cos method is a relatively recently found method for calculation of residual stress than sin method and unlike the latter, the cos method can measure the stress in single irradiation. Since the residual stress leads to Debye ring distortion, the radius at different angles is also different. The cos method measures the stress using the strain by

$$\varepsilon_{\propto} = \frac{1}{2}\left[\left(\varepsilon_{\propto} - \varepsilon_{\pi+\propto}\right) + \left(\varepsilon_{-\propto} - \varepsilon_{\pi-\propto}\right)\right] \tag{4}$$

$$\sigma_x = \frac{E}{(1+\nu)sin2\eta sin2\psi_0}\left(\frac{\partial\varepsilon_{\propto}}{\partial cos\propto}\right) \tag{5}$$

Where,

η – Mutual complementary angle of diffraction angle θ,

ψ – Angle between X-ray incidence and the normal to the sample,

σ_x – Longitudinal stress.

This method compares the distorted image with the perfect Debye ring and is all the more convenient in the stress calculation. 2DXRD makes the process of data collection and analysis much easier by using a single exposure technique, leading to a considerable reduction in measurement time of the residual stress. This, along with the fixed and miniaturised area detectors, X-ray sources, and data processing electronics, has led to the development of commercial portable instruments that enable the residual stress analysis to be conducted without having to dismount the workpiece from the equipment or any similar disruption [35]. In this study, the cos method has been used to calculate the residual stress using a recently developed portable device called Pulstec α-360s (Fig. 3(a)). This stress analysing device, manufactured by Pulstec Industrial Co. Ltd., irradiates the surface of the specimen (Fig. 3(b)) and uses an algorithm to go through the various test setting parameters and the data obtained from it. It has been used to automatically calculate the residual stress from the complete Debye ring measuring simultaneously from up to

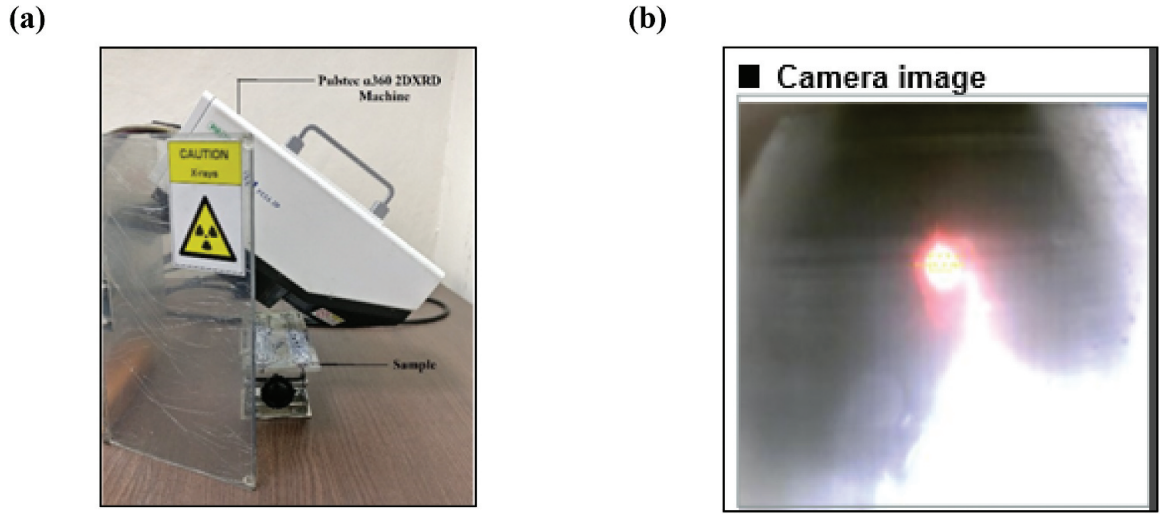


Figure 3. **(a)** Pulstec machine; **(b)** camera image of testing.

500 data points on it [36]. The X-ray signals and the distortion of the Debye ring are automatically collected and saved in the computer [37].

3. Results and discussion

(d) S4

The Pulstec stress analyser provides one with detailed analysis of all four samples, each consisting of a 3D Debye ring, 2D Debye ring and sigma(x)cos(alpha) graph (Fig. 5). The analysis also tells us about the Debye ring when opened. The sigma(x)cos(alpha) graph shows us how the curve of sigma varies with cos alpha and the gradient of the red line corresponds to the residual stress value (by cos alpha method). One can also observe across all these samples the slope of the sigma(x)cos(alpha) graph going upward indicating the presence of compressive residual stress.

Distortion rings show us the diffraction peak position and full width at half maximum (FWHM) position. The diffraction ring (Debye ring) is exactly the cycle ring in which the radius at any angle α is the same. In case of any residual stress, the Debye ring shows distortion due to an inconsistent diffraction angle in the ring, which can further help in determining the residual stress of the material [37]. During the analysis, the Debye rings are opened up in a linear form of presentation. These Debye rings are also used to analyse microstructural features such as grain orientation, texture, grain sizes, and stress states. Further, the distortion in the Debye rings tells one about any damage in the workpiece and high residual stresses [38]. In general, the FWHM of two values are related to the grain size of the film, implying that a smaller FWHM means larger the grain size and

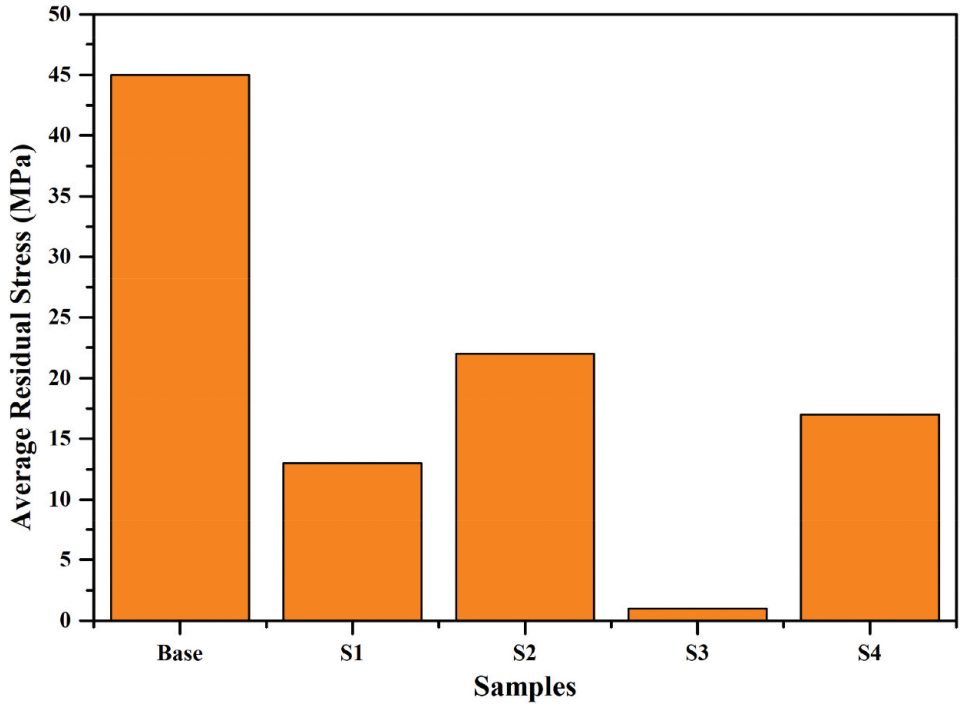


Figure 4. Experimental values of average residual stress.

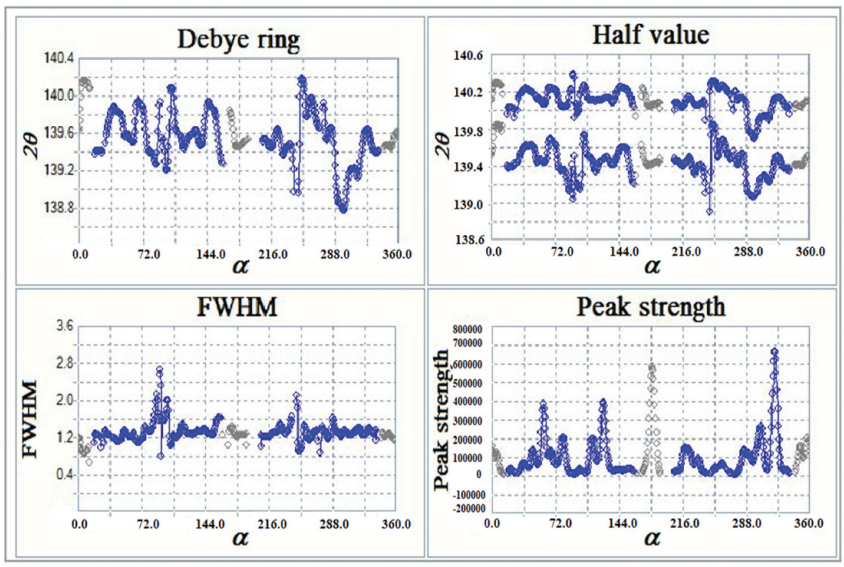


Figure 5. **(a)** Residual stress graphs of BM; **(b)** Debye ring (2D); **(c)** Debye ring (3D); **(d)** Sigma(x) cos(Alpha) graph of the base alloy.

better the crystal quality of the film [39]. FWHM is also used to determine the crystalline size of annealed materials as well as the sharpness or quality of the texture components [40,41].

Across these four samples, the compressive stress was maximum for S2 at two passes (Figure 4). It was further observed that the sample with three passes (S3) had the lowest values of both peak strength and FWHM at 1000 rpm (Figure 6), which could indicate greater grain size and lower hardness – FWHM is a measure of surface integrity [38,39].

d)

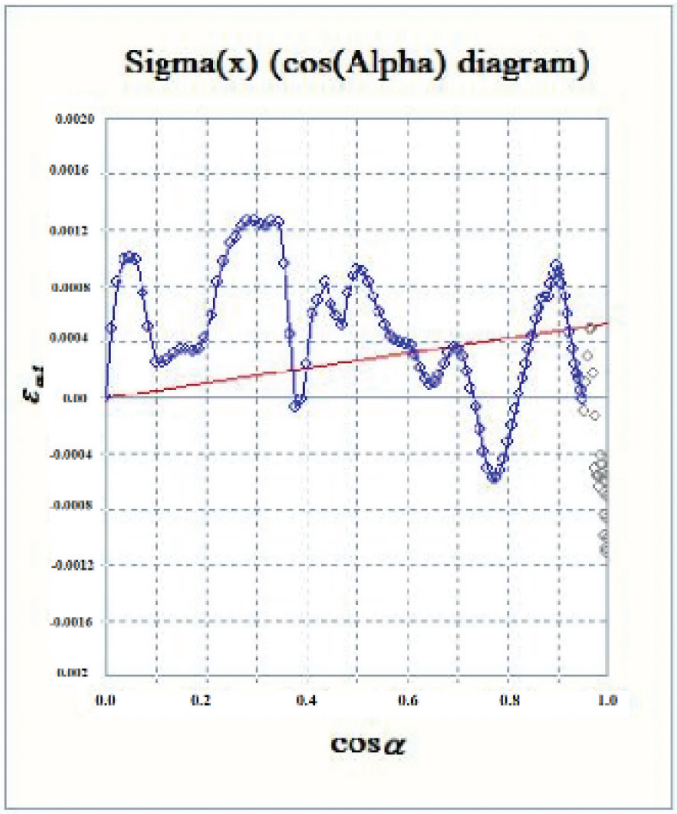


Figure 5. Continued.

As we increased the number of passes from 1 to 2, the residual stress also increased. Further increase of the number of passes leads to a decrease in residual stress and finally, again increasing the number of passes to 4, resulted in an increase in residual stress. Thus, it can be established that an even number of passes results in much more residual stress compared to an odd number of passes. The residual stresses were lowest for sample at 1000 rpm and three passes (S3).

The number of passes during FSP have a significant consequence on the surface integrity of the nanocomposites [42]. Węglowski et al. [24] showed that increase in number of passes generated higher levels of residual stress due to homogenous distribution of particles. This homogenisation results in increase in hardness and hence hardness too increases with number of passes [43–45]. However, sometimes increase in number of passes increases the heat input and causes coagulation and precipitation, and consequently a drop in hardness [46]. This is seen in S3, corroborated by low FWHM, since FWHM and hardness share a linear relationship [47,48]. Even though homogenisation was supposed to increase residual stress levels, we observe lowest residual stress in S3. This can only be attributed to the fact that increase in number of passes changed the orientation of grains, and hence consequently lowered FWHM and residual stress [49].

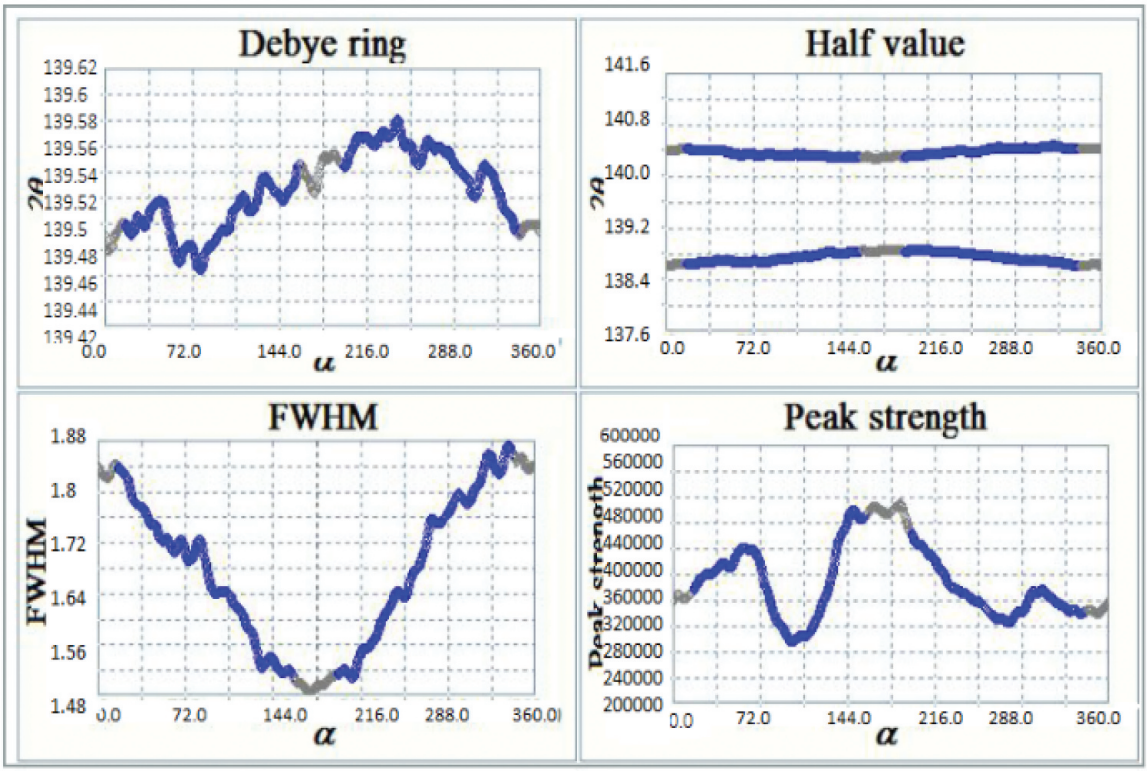


Figure 6. Debye ring, half value, FWHM and peak strength graphs of **(a)** S1; **(b)** S2; **(c)** S3 and.

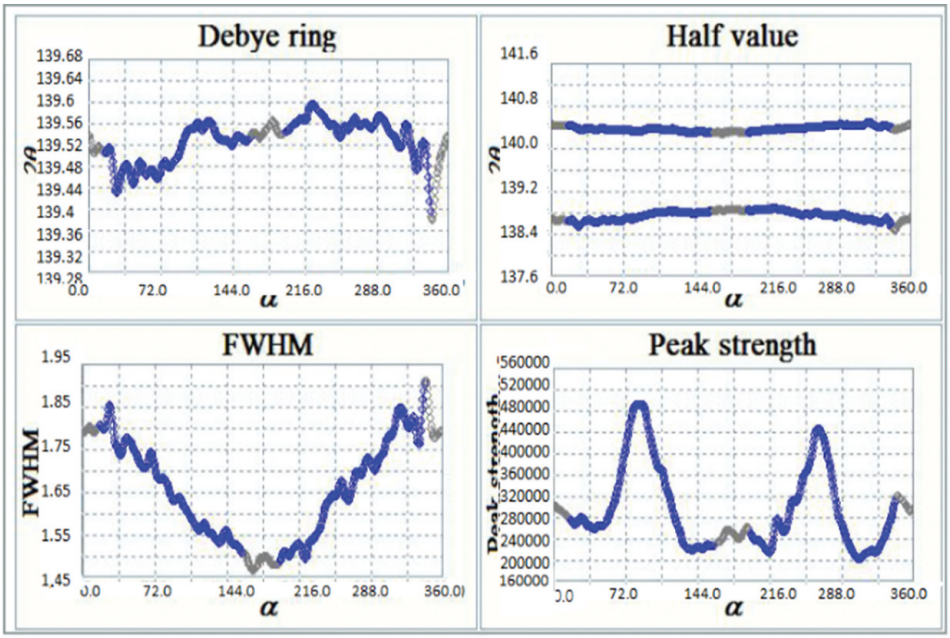
This is corroborated by the above results also. All of these parameters together have resulted in observations made above. The base alloy AA6063 was compared with the FSPed samples. The observations shown in Table 3 help us in concluding how all the four samples had a substantially lower value of compressive residual stress than the base material. The base metal also has the lowest value of FWHM compared to the four samples irrespective of the number of passes.

4. Conclusion

The aim of this study was to calculate the average residual stress values of friction stir processed surface composites of AA6063/B_4C with regards to the number of passes. It was concluded that -

- Surface composites were successfully fabricated via friction stir processing.
- Upon FSP, the residual stresses in the processed specimens were lesser than the base alloy.
- The reduction in residual stress could be attributed to better distribution of particles, evident in the residual stress graphs of the specimens.

b) S2



c) S3

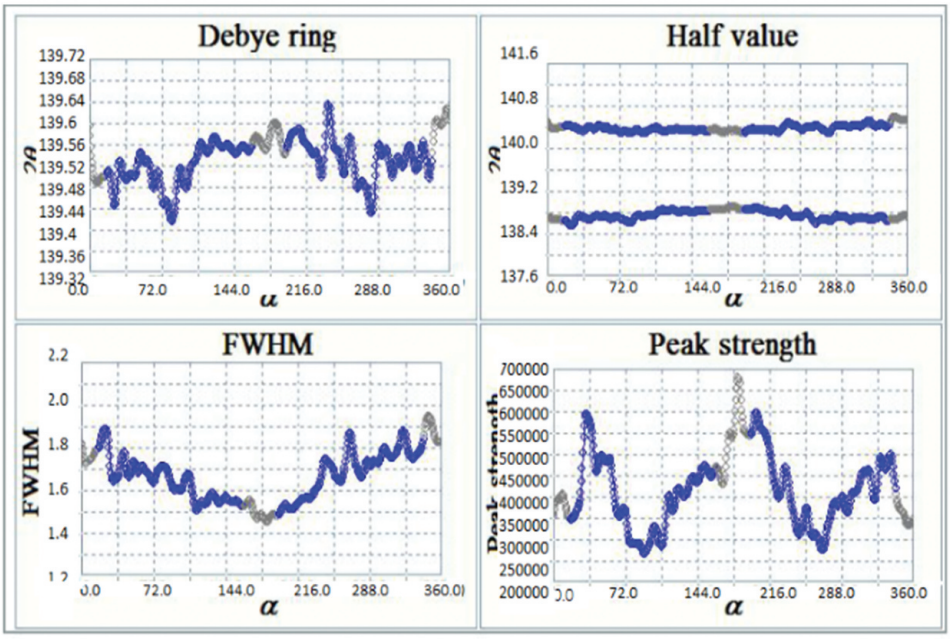


Figure 6. Continued.

d) S4

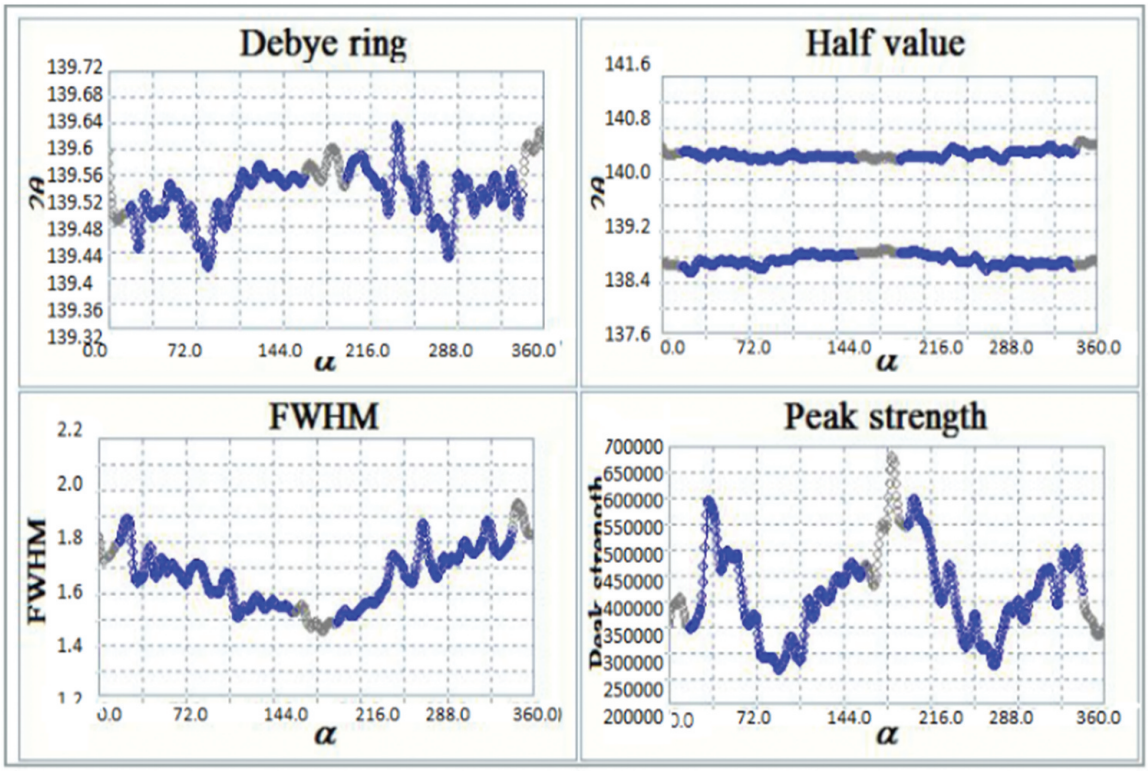


Figure 6. Continued.

Table 3. Experimental values of average residual stress of all the specimens.

Sample number	Number of passes	Tool rotational speed	Average residual stress (MPa) (compressive)
S1	1	1000	13
S2	2	1000	22
S3	3	1000	1
S4	4	1000	17
Base (AA6063)	-	-	45

- While neither of the two process parameters showed any linear relation, the most desirable residual stress characteristics were found for the specimen processed at 1000 rpm and three passes.

Disclosure of potential conflicts of interest

No potential conflict of interest was reported by the author(s).

Funding

This work was not funded.

ORCID

Kartikeya Bector http://orcid.org/0000-0002-8158-9333
Ravi Butola http://orcid.org/0000-0002-6581-4131

Data availability statement

The data used in this work may be accessed by contacting the corresponding author on the aforementioned email ID.

References

[1] Mishra R, Ma Z, Charit I. Friction stir processing: a novel technique for fabrication of surface composite. Mater Sci Eng A. 2003 Jan;341(1–2):307–310. .
[2] Rajendran C, Srinivasan K, Balasubramanian V, et al. Evaluation of load-carrying capabilities of friction stir welded, TIG welded and riveted joints of AA2014-T6 aluminium alloy. Aircr Eng Aerosp Technol. AEAT-11-2017-0233. Jun. 2019. 10.1108/AEAT-11-2017-0233.
[3] Kurtyka P, Rylko N, Tokarski T, et al. Cast aluminium matrix composites modified with using FSP process – changing of the structure and mechanical properties. Compos Struct. 2015Dec;133:959–967.
[4] Narimani M, Lotfi B, Sadeghian Z. Investigating the microstructure and mechanical properties of Al-TiB2 composite fabricated by Friction Stir Processing (FSP). Mater Sci Eng A. 2016Sep;673:436–442.
[5] Chaudhary A, Kumar Dev A, Goel A, et al. The Mechanical Properties of Different alloys in friction stir processing: a Review. Mater Today Proc. 2018;5(2):5553–5562.
[6] Butola R, Singari RM, Bandhu A, et al. Characteristics and Properties of Different Reinforcements in Hybrid Aluminium Composites: a Review. Int J Adv Prod Ind Eng IJAPIE-SI-MM. 2017. vol. 511, pp. 71–80, [Online]. Available: www.ijapie.org .
[7] Butola R, Malhotra A, Yadav M, et al. Experimental Studies on Mechanical Properties of Metal Matrix Composites Reinforced with Natural Fibres Ashes. SAE Technical Papers. 2019 Apr;2019-April:1–11.
[8] Tyagi L, Butola R, Jha AK. Mechanical and tribological properties of AA7075-T6 metal matrix composite reinforced with ceramic particles and aloevera ash via Friction stir processing. Mater Res Express. 2020;7(6). 10.1088/2053-1591/ab9c5e
[9] Khayyamin D, Mostafapour A, Keshmiri R. The effect of process parameters on microstructural characteristics of AZ91/SiO2 composite fabricated by FSP. Mater Sci Eng A. 2013Jan;559:217–221.
[10] Soleymani S, Abdollah-zadeh A, Alidokht SA. Microstructural and tribological properties of Al5083 based surface hybrid composite produced by friction stir processing. Wear. 2012Mar;278–279:41–47.
[11] Shojaeefard MH, Akbari M, Asadi P, et al. The effect of reinforcement type on the microstructure, mechanical properties, and wear resistance of A356 matrix composites produced by FSP. Int J Adv Manuf Technol. 2017 Jul;91(1–4):1391–1407. .
[12] Rajendran C, Srinivasan K, Balasubramanian V, et al. Identifying the combination of friction stir welding parameters to attain maximum strength of AA2014-T6 aluminum alloy joints. Adv Mater Process Technol. 2018 Jan;4(1):100–119. .

[13] Rajendran C, Srinivasan K, Balasubramanian V, et al. Effect of tool tilt angle on strength and microstructural characteristics of friction stir welded lap joints of AA2014-T6 aluminum alloy. Trans Nonferrous Met Soc China. 2019 Sep;29(9):1824–1835. .
[14] Butola R, Murtaza Q, Singari RM. Formation of Self-Assembled Monolayer and Characterization of AA7075-T6/B 4 C Nano-ceramic surface composite using Friction Stir Processing. Surf Topogr Metrol Prop. 2020 Jun;8(2):025030.
[15] Butola R, Singari RM, Murtaza Q. Mechanical and wear behaviour of Friction stir processed surface composite through Self-Assembled Monolayer Technique. Surf Topogr Metrol Prop. 2020;8(4):45007.
[16] Faraji G, Dastani O, Mousavi SAAA. Effect of Process Parameters on Microstructure and Micro-hardness of AZ91/Al2O3 Surface Composite Produced by FSP. J Mater Eng Perform. 2011 Dec;20(9):1583–1590. .
[17] Asadi P, Givi MKB, Abrinia K, et al. Effects of SiC Particle Size and Process Parameters on the Microstructure and Hardness of AZ91/SiC Composite Layer Fabricated by FSP. J Mater Eng Perform. 2011 Dec;20(9):1554–1562. .
[18] Rathee S, Maheshwari S, Noor Siddiquee A, et al. Process parameters optimization for enhanced microhardness of AA 6061/SiC surface composites fabricated via Friction Stir Processing (FSP). Mater Today Proc. 2016;3(10):4151–4156.
[19] Elangovan K, Balasubramanian V. Influences of tool pin profile and tool shoulder diameter on the formation of friction stir processing zone in AA6061 aluminium alloy. Mater Des. 2008 Jan;29(2):362–373. .
[20] Bector K, Tripathi A, Pandey D, et al. A Review on the Fabrication of Surface Composites via Friction Stir Processing and Its Modeling Using ANN. In: Singari RM, Mathiyazhagan K, Kumar H, editors. i*Advances in Manufacturing and Industrial Engineering*.. Singapore: Lecture Notes in Mechanical Engineering, Springer; 2021. p. 1–11.
[21] Mehdi H, Mishra RS. An experimental analysis and optimization of process parameters of AA6061 and AA7075 welded joint by TIG+FSP welding using RSM. Adv Mater Process Technol. 2020 Oct;1–23. 10.1080/2374068X.2020.1829952
[22] Tutum CC, Hattel JH. Optimisation of process parameters in friction stir welding based on residual stress analysis: a feasibility study. Sci Technol Weld Join. 2010 Jul;15(5):369–377. .
[23] Deplus K, Simar A, Van Haver W, et al. Residual stresses in aluminium alloy friction stir welds. Int J Adv Manuf Technol. 2011 Sep;56(5–8):493–504. .
[24] Węglowski MS, Sedek P, Hamilton C. "Experimental and Numerical Analysis of Residual Stress in Cast Aluminum Alloy after. Process," FSP. Key Eng Mater. 651–653:1563–1568. Jul. 2015. .
[25] Woo W, Choo H, Prime MB, et al. Microstructure, texture and residual stress in a friction-stir-processed AZ31B magnesium alloy. Acta Mater. 2008 May;56(8):1701–1711. .
[26] Jafari H, Mansouri H, Honarpisheh M. Investigation of residual stress distribution of dissimilar Al-7075-T6 and Al-6061-T6 in the friction stir welding process strengthened with SiO2 nanoparticles. J Manuf Process. 2019Jul;43:145–153.
[27] Butola R., Tyagi L., Kem L., Ranganath M.S., Murtaza Q. Mechanical and Wear Properties of Aluminium Alloy Composites:A Review. In: Sharma V., Dixit U., Sørby K., Bhardwaj A., Trehan R. (eds) Manufacturing Engineering. Lecture Notes on Multidisciplinary Industrial Engineering. 2020:369–391. Springer, Singapore. https://doi.org/10.1007/978-981-15-4619-8_28 .
[28] Dwivedi DK. Sliding temperature and wear behaviour of cast Al–Si–Mg alloys. Mater Sci Eng A. 2004 Sep;382(1–2):328–334. .
[29] Chen X-G, Da Silva M, Gougeon P, et al. Microstructure and mechanical properties of friction stir welded AA6063–B4C metal matrix composites. Mater Sci Eng A. 2009 Aug;518 (1–2):174–184. .
[30] Raaft M, Mahmoud TS, Zakaria HM, et al. Microstructural, mechanical and wear behavior of A390/graphite and A390/Al2O3 surface composites fabricated using FSP. Mater Sci Eng A. 2011 Jul;528(18):5741–5746. .

[31] Ghasemi-Kahrizsangi A, Kashani-Bozorg SF, Moshref-Javadi M, et al. Friction Stir Processing of Mild Steel/Al2O3 Nanocomposite: modeling and Experimental Studies. Metallogr Microstruct Anal. 2015 Apr;4(2):122–130.
[32] Pakkanen J, Vallant R, Kičin M. Experimental investigation and numerical simulation of resistance spot welding for residual stress evaluation of DP1000 steel. Weld World. 2016 May;60(3):393–402. .
[33] Ao S, Li C, Huang Y, et al. Determination of residual stress in resistance spot-welded joint by a novel X-ray diffraction. Measurement. 2020Sep;161:107892.
[34] Zhu, P., Gou, G., Li, Z., Zhu, M., Zhu, Z., Ma, C., & Gao, W. (2019). Study of residual stresses in A7N01 aluminum alloy with X-ray diffraction Debye ring analysis. International Journal of Modern Physics B, 33(01n03), 1940032. https://doi.org/10.1142/S0217979219400320 .
[35] Shukla S. Rapid in-line residual stress analysis from a portable two-dimensional X-ray diffractometer. Measurement. 2020Jun;157:107672.
[36] Reghu V.R., Ghosh S., Prabhu A., Shankar V., Ramaswamy P. Residual Stresses Analysis on Thermal Barrier Coatings—NDT Tool for Condition Assessment. In: Singh S., Prakash C., Ramakrishna S., Krolczyk G. (eds) Advances in Materials Processing. Lecture Notes in Mechanical Engineering. 2020;175–183. Springer, Singapore. https://doi.org/10.1007/978-981-15-4748-5_18
[37] Lin J, Ma N, Lei Y, et al. Measurement of residual stress in arc welded lap joints by cosα X-ray diffraction method. J Mater Process Technol. 2017May;243:387–394.
[38] Kumar S, Grover S, Walia RS. Effect of hybrid wire EDM conditions on generation of residual stresses in machining of HCHCr D2 tool steel under ultrasonic vibration. Int J Interact Des Manuf. 2018 Aug;12(3):1119–1137.
[39] Elilarassi R, Chandrasekaran G. Effect of annealing on structural and optical properties of zinc oxide films. Mater Chem Phys. 2010 May;121(1–2):378–384. .
[40] Patterson AL. The Scherrer Formula for X-Ray Particle Size Determination. Phys Rev. 1939 Nov;56(10):978–982. .
[41] Nittala K, Mhin S, Jones JL, et al. In situ x-ray diffraction of solution-derived ferroelectric thin films for quantitative phase and texture evolution measurement. J Appl Phys. 2012 Nov;112(10):104109. .
[42] Molla Ramezani N, Davoodi B, Aberoumand M, et al. Assessment of tool wear and mechanical properties of Al 7075 nanocomposite in friction stir processing (FSP). J Brazilian Soc Mech Sci Eng. 2019 Apr;41(4):182. .
[43] Mahmoud ERI, Takahashi M, Shibayanagi T, et al. Wear characteristics of surface-hybrid-MMCs layer fabricated on aluminum plate by friction stir processing. Wear. 2010 Mar;268 (9–10):1111–1121. .
[44] Tonelli, L., Morri, A., Toschi, S. et al. Effect of FSP parameters and tool geometry on microstructure, hardness, and wear properties of AA7075 with and without reinforcing B4C ceramic particles. Int J Adv Manuf Technol102, 3945–3961 (2019). https://doi.org/10.1007/s00170-019-03442-6 .
[45] Asadi P, Faraji G, Besharati MK. Producing of AZ91/SiC composite by friction stir processing (FSP). Int J Adv Manuf Technol. 2010 Nov;51(1–4):247–260. .
[46] Gangil N, Maheshwari S, Siddiquee AN. Multipass FSP on AA6063-T6 Al: strategy to fabricate surface composites. Mater Manuf Process. 2018 May;33(7):805–811. .
[47] Rai S, Choudhary B, Jayakumar T, et al. Characterization of low cycle fatigue damage in 9Cr–1Mo ferritic steel using X-ray diffraction technique. Int J Press Vessel Pip. 1999 Apr;76 (5):275–281.
[48] Pariente I, Guagliano M. About the role of residual stresses and surface work hardening on fatigue ΔKth of a nitrided and shot peened low-alloy steel. Surf Coat Technol. 2008;202:3072–3080.
[49] Ongtrakulkij G, Khantachawana A, Kondoh K. Effects of media parameters on enhance ability of hardness and residual stress of Ti6Al4V by fine shot peening. Surf Interfaces. 2020Mar;18:100424.

Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021)
IEEE Xplore Part Number: CFP21K74-ART; ISBN: 978-0-7381-1327-2

Study of robust facial recognition under occlusion using different techniques

Deepanshu Jayant
Master of Technology
Department of Computer Science
Delhi Technological University, New Delhi
deepanshujayant@gmail.com

Dr. Aruna Bhat
Associate Professor
Department of Computer Science
Delhi Technological University, New Delhi
aruna.bhat@dtu.ac.in

2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) | 978-1-6654-1272-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICCS51141.2021.9432300

***Abstract*— This paper represents the different techniques used for the study of robust facial recognition. The need of facial recognition is expanding very rapidly in current technologies as it is possible to identify a humane face features through different digital mediums. Although, facial recognition is not done achieved easily in the real-world situation as most of the time the human face goes through various occlusions making hard for the system to complete the identification process. The sole purpose of this article is to compare the techniques used and identify the most accurate and efficient one from all of them. Robust facial recognition is a task of doing recognition under uncertain environments and features. The main aim of the robust facial recognition is to analyze and identify the face images where the picture proposes difficult viewpoints, angles, poses, illumination, noises and expressions. The techniques studied for this article are Local Binary Patterns, HOG features, Occlusion-adaptive Deep Networks (ODN), Robust Principal Component Analysis (RPCA), Progressive Convolutional Neural Network and Region Attention Networks (RAN). In this article, a table is created summarizing the experiments and results of the techniques studied.**

Keywords—LBP, HOG, ODN, RPCA, PCA, RAN, CNN

I. INTRODUCTION

Recognition of facial expressions involves understanding changes in facial features such as eyes, eyebrows etc. In general, changes in the shape and shape of the groove or wrinkles are outstanding features that can be used for facial analysis. And facial expressions are also one of the most powerful, natural, quick ways for people to talk about their feelings and intentions. Because of its potential applications, automatic facial recognition has attracted a lot of attention over two decades. Facial recognition methods often use face geometry and visual features to distinguish different expressions. The shape of facial objects and their related changes are arranged under geometric features. The different technique such as Robust Principal Component Analysis (RPCA) [2], Local Binary Patterns (LBP) [1], Histogram of Oriented Gradients (HOG) [3], Principal Component Analysis (PCA), Local Gabor Binary Patterns [20], Gabor Descriptor [19], etc., encodes the surface texture of the facial regions. The movement of facial features reveals wrinkles or folds of skin that are effectively delivered by the visual elements. Although, the face in the real-world datasets often consists of occlusions, which sometimes covers the important facial features. Hence, making it difficult to identify the human face. Therefore, this article focuses on the techniques such as HOG, LBP, RPCA, ODN and RAN.

II. METHODOLOGIES

1. Local Binary Patterns

1.1. Definition

Local binary patterns can be defined as simple operator which turn out to be very effective and efficient texture operator. LBP works by labeling the pixels of a picture by setting up a threshold for the neighboring pixels and calculates the results in the form of binary numbers. Being first introduced in 1994 by Ojala et al, it is still found to be very effective feature for the facial recognition.

1.2. Working

The Local Binary Patterns algorithm uses 4 parameters to produce an outcome. The parameters are (i) Radius, (ii) Neighbors, (iii) Grid X and (iv) Grid Y.
Where,
Radius is used as an entity to build the circular local binary patterns and use it to represent the radius of the central pixels. Most of the time value of radius is set to 1.

Neighbors are defined as the sample points which are used in the building of circular local binary pattern. Most of the time value is set to 8.

Grid Y and **Grid X** are the number of blocks present in the vertical and horizontal directions respectively. The greater number of cells represents the clarity level of grids in each direction. Its value is mostly set to 8.

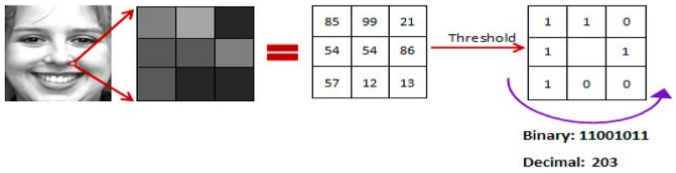


Fig. 1 Local Binary Pattern Procedure

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:43:03 UTC from IEEE Xplore. Restrictions apply.

1.3 Application

The local binary pattern algorithm is trained by using a dataset with the several facial images of the people. Also, for accurate results each image is indexed with a unique ID so that algorithm can use the information to recognize the input image to produce accurate outputs.

The algorithm is applied to a window, in which a neighborhood circle is defined. The output is produced by comparing the central pixel with another 8 neighboring pixels.

Formula used:-

$$LBP_{P,R}(x,y)=\sum_{p=0}^{P-1} s((f_{xp}-y_p)-f(x,y))2^p$$

$$s(z)=\{1,\ z\geq 0 \quad 0,\ z<0 \qquad (1)$$

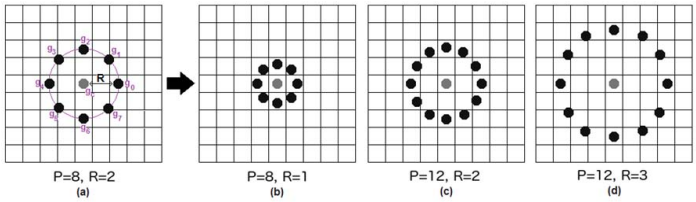


Fig.2. Showing comparison performed with neighboring pixels

The result obtained afterwards is an 8-bit binary string which helps in the creation of histogram of the feature vector.

2. Histogram of Oriented Gradient

2.1 Definition

The Histogram of Oriented Gradient (HOG) is a framework used mainly for the detection of faces in the number of images and pictures. While HOG is only a decade old, it offers numerous applications such as autonomous vehicles, surveillance applications, etc. Also, its growth is increasing very rapidly than other method because of its efficient results. To produce an effective outcome HOG algorithm uses feature descriptor, pre-processors and gradients.

2.2 Working

HOG algorithm has two preliminary steps to calculate the results of the facial recognitions. Which are stated as follows,

Feature Descriptors are used to extract the necessary information from the scanned image and helps in discarding any unwanted information from the sample image. Though, the produces image from descriptors is hard to look at but gives a perfect window for other algorithms to produce good results.

Pre-processing is used to fix the aspect ratio of the sample imaged used which allows the algorithm for easier accessibility. A common ratio is set as 1:2, for example, 100x200 or 500x1000.

2.3 Application

The HOG feature descriptor requires a histogram which is used in the algorithm for the desired results. The histogram is created by calculating the respective horizontal and vertical gradients. Gradients are calculated by simply filtering the images using kernels,

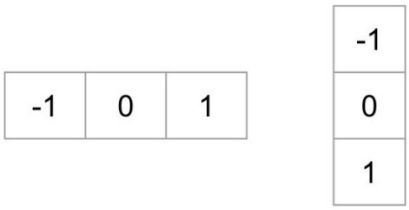


Fig.3. Kernels

The magnitude and the respective directions of the gradients are found using the following formulas,

$$g = \sqrt{g_x^2 + g_y^2}$$

$$\theta = \arctan \frac{g_y}{g_x} \qquad (2)$$

Final step includes the creation of the histogram using the calculated gradients, which is done by dividing the images into cells according to the resolution of the sample image. In this article, 4x4 cells are studied for the HOG methodology.

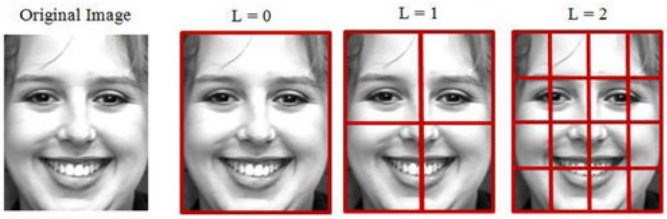


Fig. 4 HOG procedure of defining cells

Therefore, histogram vector is calculated with respect to the size and direction values in each cell present in the matrix created for the sample image.

3. Occlusion-adaptive Deep Networks

3.1 Definition

Occlusion-adaptive Deep Networks is a simple and an efficient methodology proposed to overcome the irregularity occurs due to the occlusions in the sample images for the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:43:03 UTC from IEEE Xplore. Restrictions apply.

facial recognition. The Occlusion-adaptive Deep Networks includes three terminologies such as (i) geometry-aware module, (ii) distillation module and (iii) low-rank learning module [5].

3.2 Working

Distillation Module is used to model the occlusion present in the image. It is used to map the occlusion potential in terms of high-quality features, whose main purpose is to serve as the adaptive weight to reduce the levels of occlusion in the sample image.

Low-rank Learning Module is used to collect the complete features of the face as reduction of occlusion levels disrupts the due to the absence of original semantic features of the face. It is done through shared structural matrix.

Geometry-aware Module is an important method as it helps the low-rank learning module to recollect the absent semantic from the image using the geometric data of the face such as position, relation of the elements present in the face.

3.3 Application

ODN works by modifying the residual unit of ResNet-18 [18] to the working ODN model. The method works by feeding the feature map collected from the residual to the geometry-aware module and the distillation module at the same time, which enables the framework to map the geometrical features of the face present in the sample image.

Hence, the outcomes produced by both modules are then fed to the third module, i.e. low-rank learning module, which recovers the missing features of the face lost due to the present occlusion by modeling the inner-features of face. The working and relationship among the modules is described in the fig.5. below,

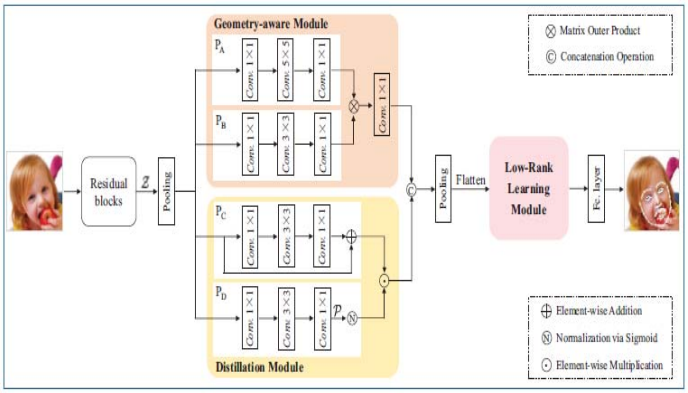


Fig. 5 Working and Relationships of modules

4. Robust Principal Component Analysis

4.1 Definition

PCA is a well-known size the most commonly used reduction method in the area of machine learning and computer sight. Used for converting peak-quality data to a low-level location and widely available for its uses in face recognition. Since, PCA is an influential mathematical tool, the most common PCA at risk of noisy data and even a single corruption entry can change the magnitude of the rating. To solve the major problem facial images, suffer such as occlusions, noise or illumination variations. The Robust PCA is proposed, and the work presented by Candes and Emmanuel J [16] is focused in this article.

The RPCA is a way of saying that reduces the data by breaking data matrix into two parts, low-position and sparse. Low-position matrix restores crucial features and structures of facial expression while a sparse-matrix stores the noisy data.

4.2 Application

As RPCA uses data matrixes for the decomposition into the different components. The Principal Component Pursuit (PCP) [16] problem has been observed. As, there are several ways to solve this issue, the Augmented Lagrange Multiplier (ALM) [17] algorithm is applied for the same. Therefore, when applied with RPCA for facial recognition, the primary data matrix, says Z, is decomposed into low ranked matrix, say X, and a sparse-matrix, say Y, as shown in the fig. Here, the image shown renders low level of sparse noise, hence producing a better representation of the sample image via matrix X than the original data matrix Z.

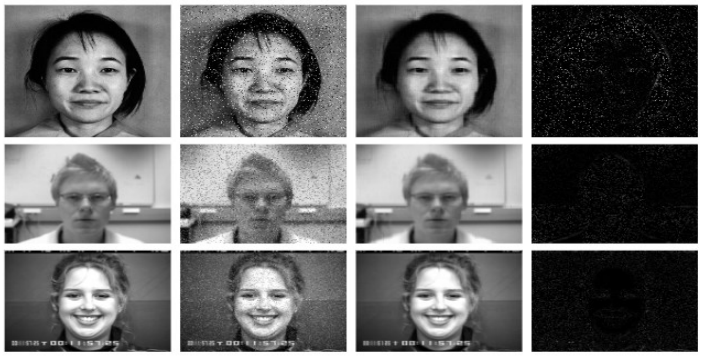


Fig. 6 Results obtained after using RPCA

Now, the primary PCA can be easily performed as resulted image is not of high dimension. Hence, the PCA now is applied to matrix A rather than matrix X using the following formula,

$$\frac{\sum_{i=1}^{k}\lambda_i}{\sum_{i=1}^{p}\lambda_i} >= 0.95 \quad (3)$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:43:03 UTC from IEEE Xplore. Restrictions apply.

5. Region Attention Networks

5.1 Definition

Region Attention Networks (RAN) works by capturing the important elements of facial regions for occlusion and pose adaptively. It embeds and aggregates several values of region features calculated by CNN into a compact representation. RAN provides region biased loss to divert the attention towards the important facial regions.

Region Attention Networks includes three modules to calculate the desired outcome, which are (i) feature extraction module, (ii) self-attention module and (iii) relation attention module. [4]

5.2 Working

RAN works by studying the most targeted weights for each key region in the end and combines CNN features into a combined presentation. Main purpose of the modules is to learn the weights of the environment and to refine it according to the global context. Also, the framework exhibits two supplementary effects on the sample facial images.

Region Attention Network proposes a region biased loss [4] system (RB-Loss) which encourages an attention weight to the crucial facial features. The RB-Loss limits the RAN model so that the maximum attention weight of the face areas should be greater than the original image of the sample face.

The region feature set is defined by,

$$X = [F_0, F_1, \cdots, F_k] = [r(I_0;\theta), r(I_1;\theta), \cdots, r(I_k;\theta)], \tag{4}$$

5.3 Application

Self-Attention Module [4] uses the FC layer and sigmoid function to measure the weights of the most attention to the regional features. The weight of attention is defined by the following formula,

$$\mu_i = f(F_i^\top \mathbf{q}^0), \tag{5}$$

As the module summarizes the region features into a global representation using attention weights calculated, compact represented is shown as,

$$F_m = \frac{1}{\sum_{i=0}^{n} \mu_i} \sum_{i=0}^{n} \mu_i F_i. \tag{6}$$

Relational-Attention Module aim is to refine the attention weights further by modelling the relation between the features and global representation. The new attention weights are calculated by the sample concatenation and another FC layer and is formulated as,

$$v_i = f([F_i : F_m]^\top \mathbf{q}^1), \tag{7}$$

After aggregating all the regional data alongside global representation produces by self-attention module, the new compact feature is formulated as,

$$P_{RAN} = \frac{1}{\sum_{i=0}^{n} \mu_i v_i} \sum_{i=0}^{n} \mu_i v_i [F_i : F_m]. \tag{8}$$

Region Biased Loss is a measure to identify the crucial facial features which requires high attention weights for the efficient outcome. It defines a constraint enforcing that the attention weights must be larger than the original image with a margin.

RB-Loss is described as,

$$\mathcal{L}_{RB} = \max\{0, \alpha - (\mu_{max} - \mu_0)\}, \tag{9}$$

The relationship and flow of outcome produced by the model is shown in the Fig below,

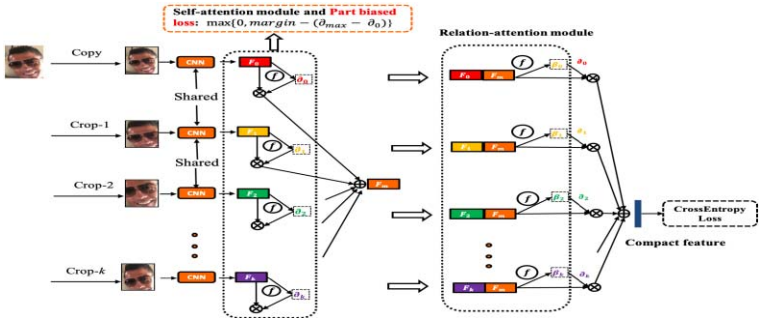


Fig.7. RAN flow of outcome

III. EXPERIMENTS

This section will showcase the experimentation, datasets and the results of the methodologies studied in this article.

As discussed earlier, this article focused on five different methodologies, i.e. LBP, HOG, RPCA, RAN and ODN. All methods have implemented the experiments on different sets of datasets of facial images.

A. Datasets

Local Binary Patterns and **Histogram of Oriented Gradients** framework is train and tested on Cohn-Kanade (CK+) Facial Expression Database [13].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:43:03 UTC from IEEE Xplore. Restrictions apply.

Robust Principal Component Analysis framework is train and tested CK+, Japanese Female Facial Expression (JAFFE) database [15] and Oulu CASIA database [14].

Occlusion-adaptive Deep Networks framework is train and tested on three challenging datasets, including 300W [8], COFW [6] and AFLW [7].

Region Attention Networks framework is train and tested on challenging datasets like FERPlus [9], AffectNet [12], RAF-DB [11] and SFEW [10].

Author	Framework	Algorithms	Performance	Database	Limitations
Caifeng Shan, Shaogang and Peter W. McOwan [1]	Local Binary Patterns	LBP + Template Matching LBP + SVM	79.1 88.4 (Polynomial)	CK+ CK+	Performs well only when low-res video input is available
Pranav Kumar, S L Happy and Aurobinda Routray [3]	Histogram of Oriented Gradients	Linear SVM RBF SVM	89.0 89.0	CK+ CK+	Speed and recognition accuracy purely depends on the dynamics of expressions
Saloni Mundra, Sujata and Suman K. Mitra [2]	Robust Principal Component Analysis	-	84.93 (k=11) 91.67 (k=25) 99.16 (k = 18)	JAFFE CK+ Oulu CASIA	While achieving good accuracy with noisy data, it may differ when pose dataset is used.
Kai Wang, Xiaojiang Peng, Jianfei Yang, Debin Meng and Yu Qiao [4]	Region Attention Networks	RAN (w RB-Loss) RAN (w RB-Loss) RAN (w RB-Loss)	58.50 82.72 83.63	AffectNet(Occlusion) RAF-DB(Occlusion) FERPlus(Occlusion)	The framework is sensitive to the calculations of Region biased loss (RB- Loss)
Meilu Zhu, Daming Shi, Mingjie Zheng, and Muhammad Sadiq [5]	Occlusion-adaptive Deep Networks	NRMSE (x 10^{-2})	6.67 1.63	300W AFLW	While achieving good results on benchmark datasets, framework is dependent on baseline of BRNet.

Table 1 Summarized data of the methodologies focused in the article

B. Results of Experimentation

The summary of the different methodologies, their methods and limitations are shown in the table shown,

IV. CONCLUSION

This article concludes that the simple yet effective frameworks such as **Local Binary Patterns** and **Histogram of Oriented Gradients** shows an efficient and great results when the image dataset is noiseless and have low levels of occlusion and pose. While **Robust Principal Component Analysis** outperforms the LBP and HOG frameworks when images used possesses noisy data in them but does not performs well the sample images have high level of occlusion and pose variations.

To overcome the invariance raised by the levels of occlusions and pose in the real-world image datasets. Frameworks such as **Region Attention Networks (RAN)** and **Occlusion-adaptive Deep Networks (ODN)** outperform the previous frameworks by an effective margin. While RAN achieves very much high accuracy results, ODN has outperformed existing methods while being tested with benchmark datasets and has achieved robustness against occlusion effectively.

REFERENCES

[1]. Caifeng Shan, Shaogang Gong and P. W. McOwan, "Robust facial expression recognition using local binary patterns," *IEEE*

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:43:03 UTC from IEEE Xplore. Restrictions apply.

International Conference on Image Processing 2005, Genova, Italy, 2005.

[2]. S. Mundra, Sujata and S. K. Mitra, "Modular Facial Expression Recognition on Noisy Data Using Robust PCA," *2019 IEEE 16th India Council International Conference (INDICON)*, Rajkot, India, 2019

[3]. P. Kumar, S. L. Happy and A. Routray, "A real-time robust facial expression recognition system using HOG features," *2016 International Conference on Computing, Analytics and Security Trends (CAST)*, Pune, India, 2016.

[4]. K. Wang, X. Peng, J. Yang, D. Meng and Y. Qiao, "Region Attention Networks for Pose and Occlusion Robust Facial Expression Recognition," in *IEEE Transactions on Image Processing*, 2020.

[5]. M. Zhu, D. Shi, M. Zheng and M. Sadiq, "Robust Facial Landmark Detection via Occlusion-Adaptive Deep Networks," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, 2019.

[6]. Xavier P. Burgos-Artizzu and Pietro Perona. Robust face landmark estimation under occlusion. In *International Conference on Computer Vision*, pages 1513–1520, 2013.

[7]. Peter M. Roth Martin Koestinger, Paul Wohlhart and Horst Bischof. Annotated Facial Landmarks in the Wild: A Largescale, Real-world Database for Facial Landmark Localization. In *IEEE International Conference on Computer Vision Workshops*, pages 2144–2151, 2011.

[8]. Christos Sagonas, Georgios Tzimiropoulos, Stefanos Zafeiriou, and Maja Pantic. 300 faces in-the-wild challenge: The first facial landmark localization challenge. In *IEEE International Conference on Computer Vision Workshops*, pages 397–403, 2014.

[9]. E. Barsoum, C. Zhang, C. C. Ferrer, and Z. Zhang, “Training deep networks for facial expression recognition with crowd-sourced label distribution,” in *Proc. 18th ACM Int. Conf. Multimodal Interact. (ICMI)*, 2016.

[10]. A.Dhall, R.Goecke, S. Lucey, and T. Gedeon, “Static facial expression analysis in tough conditions: Data, evaluation protocol and benchmark,” in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCV Workshops)*, Nov. 2011.

[11]. S. Li and W. Deng, “Reliable crowdsourcing and deep localitypreserving learning for unconstrained facial expression recognition,” *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 356–370, Jan. 2019.

[12]. A.Mollahosseini, B. Hasani, and M. H. Mahoor, “AffectNet: A database for facial expression, valence, and arousal computing in the wild,” *IEEE Trans. Affect. Comput.*, vol. 10, no. 1, pp. 18–31, Jan. 2019.

[13]. P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I. Matthews, ”The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression,” 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, San Francisco, CA, 2010, pp. 94-101.

[14]. Guoying Zhao, Xiaohua Huang, Matti Taini, Stan Z. Li, and Matti PietikInen. 2011. Facial expression recognition from near-infrared videos. Image Vision Comput. 29, 9 (August 2011), 607-619.

[15]. Michael Lyons, Shigeru Akamatsu, Miyuki Kamachi, and Jiro Gyoba.

[16]. Cands, Emmanuel J., et al.”Robust principal component analysis?” Journal of the ACM (JACM) 58.3 (2011): 11. New York: Academic, 1963, pp. 271–350.

[17]. Z. Lin, M. Chen, L. Wu, and Y. Ma. The augmented Lagrange multiplier method for exact recovery of a corrupted low-rank matrices. Mathematical Programming, submitted, 2009. Coding facial expressions with gabor wavelets. In Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on, pages 200205. IEEE, 1998.

[18]. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In IEEE Con- ference on Computer Vision and Pattern Recognition, pages 770–778, 2016.

[19]. W. Gu, C. Xiang, Y. Venkatesh, D. Huang, and H. Lin, “Facial expression recognition using radial encoding of local gabor features and classifier synthesis,” Pattern Recognition, vol. 45, no. 1, pp. 80–91, 2012.

[20]. T. R. Almaev and M. F. Valstar, “Local gabor binary patterns from three orthogonal planes for automatic facial expression recognition,” in Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on. IEEE, 2013, pp. 356–361.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on June 07,2021 at 08:43:03 UTC from IEEE Xplore. Restrictions apply.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/352246180

Super-Capacitor based Electric Vehicle Electric Vehicle Charging

Article *in* International Journal of Engineering and Technical Research · June 2021

CITATIONS
0

READS
87

4 authors, including:



Hemant Sharma
Delhi Technological University
2 PUBLICATIONS **0** CITATIONS
SEE PROFILE



Imroz Khan
Delhi Technological University
2 PUBLICATIONS **0** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Multi-output fast wireless charging for electric vehicle View project

SUPER-CAPACITOR BASED ELECTRIC VEHICLE View project

All content following this page was uploaded by Hemant Sharma on 09 June 2021.

The user has requested enhancement of the downloaded file.

Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 05, May-2021

Super-Capacitor based Electric Vehicle Electric Vehicle Charging

Hemant Sharma
Student of Electrical Engineering
Delhi Technological University
Delhi,India

Imroz Khan
Student of Electrical Engineering
Delhi Technological University
Delhi,India

Anjali Kumari
Student of Electrical Engineering
Delhi Technological university
Delhi,India

Prof. Prem Prakash
Faculty of Electrical Engineering
Delhi Technological Engineering
Delhi, India

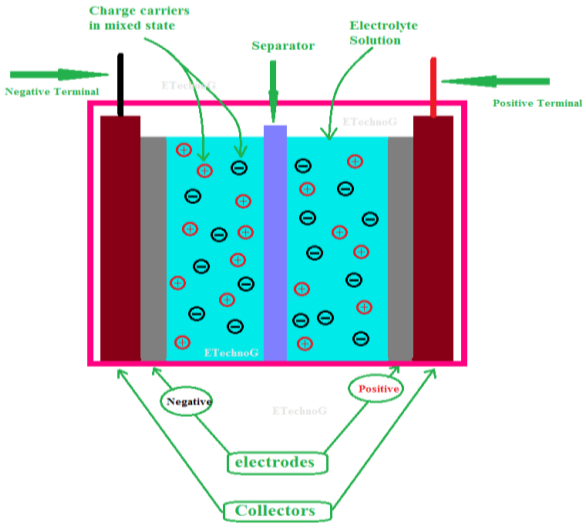
***Abstract* — It is essential to maintain a greener and safer environment by reducing carbon emissions from our earth progressively. Carbon emissions will generally be decreased by a substantial reduction in fossil fuel-powered vehicles and a transition to electric cars, which use produced electricity as fuel, emit fewer emissions than a traditional vehicle, and are zero-emission if the electricity used is renewable. Supercapacitors are widely used nowadays. These High-pressure, high-efficiency energy storage devices are also known as Ultracapacitors or electrochemical double-layer capacitors (EDLC). Their favorable properties make them ideal for use in energy storage systems, including the ability to be charged and discharged quickly without losing performance over a long period. A supercapacitor pack can be used in HESS (battery-supercapacitor system), which integrates various energy storage technologies with a specialized control strategy that maximizes the benefits of each energy source used to achieve overall performance. This research paper will describe a brief overview of the benefits, features, advantages, and disadvantages of hybrid energy systems based on batteries and supercapacitors.**

Keywords—Wireless charging system; Electric vehicle; Hess system; Supercapacitor; Hybrid vehicle.

INTRODUCTION OF WIRELESS CHARGING

In contemporary-day delivery, increasingly electric-powered vehicles (EV) of various kinds may be visible every day: electric-powered cars, electric-powered buses, electric-powered scooters, electric-powered motorcycles etc. The motives of growing the quantity of EV on the roads lie in each ecology and their performance. Thus, despite being more luxurious to buy, EVs devour much less power and are cheaper in exploitation. It is crucial to enforce excessive performing electric energy storage additives concerning their lifetime, power density, energy density, cycle performance, cost, length, and better garage overall performance. combining each component - battery and supercapacitor – to shape a hybrid energy storage system (HESS) that could grow the general overall performance of the electric cars via means of storing the energy from acceleration competencies to the deceleration of the vehicle. It is known that in conventional HESSs, the battery is directly related to the DC hyperlink, whilst in a 1/2 of bridge DC/DC converter, the battery is positioned among the supercapacitor and the DC hyperlink. The essential benefit of the supercapacitors is the potential to charge and discharge constantly without degrading and functionality for working excessive energy score in comparison to batteries. In this approach to be able to provide energy storage for electric cars, each charge sustaining and plug-in designs have to make use of supercapacitors in aggregate with batteries.

SUPERCAPACITOR STRUCTURE



In a battery-capacitor energy system, the galvanic battery serves as an electricity supply for long-distance touring at the same time as the supercapacitor pack is used as a peak power supply providing battery lifetime improvements, power for acceleration and the opportunity for complete regeneration of energy for the duration of braking which improves energy performance and usage of electric energy. The standard overall performance of supercapacitors relies on the selected material for the electrode, electrolyte, separator and contemporary collector. Supercapacitor's electrolytes need to be cautiously decided on to reduce their internal resistance.

The main components in the construction of supercapacitors are two charged electrodes, a current collector and a separator that allows for the transfer of ions and prevents direct

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 05, May-2021

electrical contact. When an electric charge is collected among the primary electrode and electrolyte, the equal quantity of charges with contrary polarity might be precipitated on the second electrode forming charged layers with minimum separated distance.
Supercapacitors are the precise solution when a quick charge is needed to provide a short-term energy. At the same time, batteries are regularly selected to offer long-time period energy as they can be recharged very quickly. Also, they are appropriate and extra robust to bridge power gaps persevering with from some seconds to a few minutes.

To compensate for the challenge discovered in battery-powered EVs, SCs are used along with the battery, ensuing in a hybrid ESS(HESS). Their excessive power density and potential characterize the SCs to work on a vast variety of temperature and their potential to respond quickly. The typical characteristics of each battery and SCs are provided. Hence hybrid ESS (HESS) topologies were proposed together with both batteries and SCs.

TYPICAL CHARACTERSTICS OF BOTH BATTERIES AND SUPER CAPACITOR

Energy Storage System	Rated Power (MW)	Typical discharge time	Energy density (W-h/kg)	Power density (W/kg)	Response time	Cycle lifetime (in cycles)	Lifetime (in years)	Efficiency (%)	Working temperature range
Li-ion Battery	0.1-100	Hours	130-250	230-340	ms	4000	5-8	65-95	-10°C to 50°C
SCs	0.01-0.3	ms-min	0.05-15	$10-10^6$	ms	-	-	85-98	-40°C to 60°C

Note. Reprinted from K. Jayasawal, A. K. Karna and K. B. Thapa, "Topologies for Interfacing Supercapacitor and Battery in Hybrid Electric Vehicle Applications: An Overview," *2021 International Conference on Sustainable Energy and Future Electric Transportation (SEFET)*, 2021.

Supercapacitors can be classified and distinguished mainly in three types depending on the cell configuration or energy storage system, electric double layer capacitors, hybrid asymmetric capacitors and pseudo capacitors.

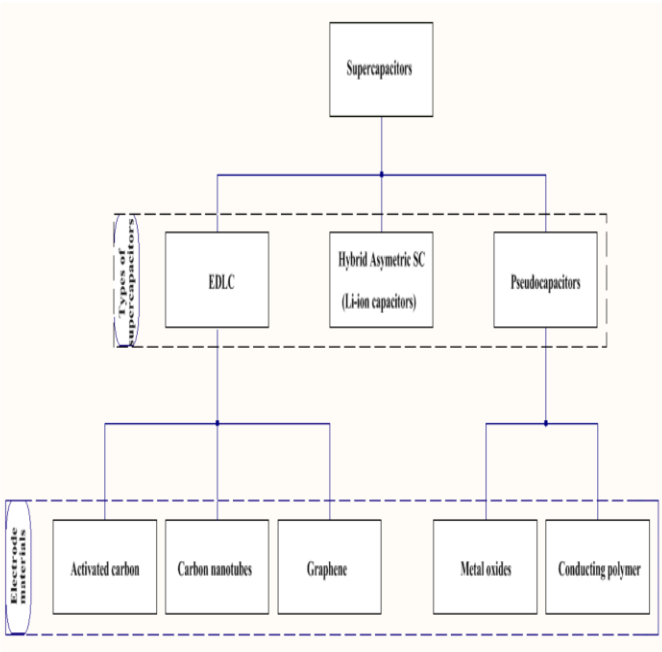


Fig. 1. CLASSIFICATION OF SUPERCAPACITOR

With relevance to EDLC capacitors, the storage of electrical energy is achieved by charge separation in Helmholtz double-layer acting as a boundary between the conductor and electrolyte. Pseudocapacitors have chemical compound conducting electrodes or metal chemical compound primarily based on electrodes combining the static and pseudo capacitance charge-storage process. Hybrid capacitors have uneven electrodes composed of a double layer capacitor electrode and a pseudocapacitive electrode, incorporating the most straightforward options of each technology. Li-ion capacitors use one electrostatic conductor and one chemistry electrode letting higher energy density and self-discharge characteristic than an EDLC capacitor, additional charge-discharge cycles than a Li-ion battery while not the potential for risky thermal runaway.

ADVANTAGE OF HAVING A HYBRID ENERGY STORAGE SYSTEM

Many topologies of HESS are best-known, like supercapacitor/battery, battery/supercapacitors, half-bridge, full-bridge and multiple input converter topologies that are investigated, developed, designed and examined over the last twenty years. As mentioned above, the combination of a battery and a supercapacitor forms a HESS.
Advantages - 1) extension of battery cycle-life, 2) growth of the overall powertrain efficiency, 3) growth of the electricity capacity of the powertrain. Well-designed HESS and its overall performance rely upon the potential to satisfy the power requirement described with the aid of using the individual overall performance parameters of the battery and supercapacitor packs and extended operational flexibility.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 05, May-2021

WIRELESS CHARGING

Wireless charging has started an alternative method for charging electric gears. As charging of battery throw wireless technology for Electric Vehicles (EVs) arise numerous issues just like the power pad, coil association plan's and most sizeable is charging time for batteries of an electric-powered car, power converters for excessive recurrence strength change, and also electromagnetic field securities. The inductive coupling systems typically operate in khz band as a result of the state-of-artwork power electronic devices are to be had for every power generation and conditioning. On the opposite hand, this low frequency needs an enormous length coil and heavy ferrite materials, which could not be most well-liked with using vehicles in terms of payload efficiency. The resonance coupling systems work higher operating frequencies within the megahertz band. it's taken into thought to be promising for the reason of the wireless car charging owing to the subsequent advantages:

1. More efficiency at less distance
2. High transmission distance with moderate efficiency
3. solely act with the resonant body (lower magnetic attraction exposure to non-resonant body)
4. very less weight (no iron and ferrite core required)
5. compact size

Wireless energy transfer suggests the electrification of cars might extend on the far side, delivering electricity and changing it into energy through the onboard batteries of stationary vehicles. Especially, the wireless charging of moving vehicles on-demand and in the period of time (i.e., dynamic charging) would result in a paradigm shift of standard transit. The wireless charging of electric vehicles can considerably alleviate the demand for onboard batteries or perhaps modify battery-free vehicles. Considering the need for quick and frequent wireless charging, another variety of electricity storage device, supercapacitors, can be additional appropriate than batteries thanks to their glorious characteristics for vehicle onboard usage:

1. Work electrostatically while not reversible chemical reactions
concerned
2. unlimited cycle life (use of cycle is unto million times)
3. quick and high economical charge/discharge thanks to little internal
resistance with about 97-98 % efficiency
4. (charged voltage is proportional to the energy stored in a capacitor)
5. An operation temperature vary from -40 to +70◦C and
a little amount of leakage current.
6. No need to use heavy material so environmentally friendly.

DIFFERENT TOPOLOGIES FOR HESS

The first topology is a partially–decoupled HESS. In this configuration, the supercapacitor is attached immediately to a bidirectional DC/DC converter. The battery is attached to the terminal of the inverter (DC/AC converter). The DC hyperlink voltage can have small voltage fluctuations because the battery is connected to the DC hyperlink. But here, the electricity density of the supercapacitor may be fully applied as it may be discharged to zero volts. The dangers of this configuration are that the DC/DC converter has to modify the strength speedy, which increases the implementation complexity.

Another HESS topology is every other partially decoupled shape. The battery is attached to a bidirectional DC/DC converter. The supercapacitor is related to the terminals of the DC hyperlink running as a low pass filter. The power that goes with the drift may be efficiently controlled. The supercapacitor can work in a much wider variety, and the DC hyperlink voltage can reveal excessive fluctuations.

DISCUSSION

The SCs, in contrast to batteries, because of their traits, are capable of offering an immediate reaction for a changing demand application, at the same time as being capable of performing in harsh environmental conditions. It, in hybridization with battery, can increase the battery's existence and extend battery substitute duration. The SCs can release and soak up excess energy and might offer the use of power electronic converter and switches, can offer bendy voltage operations at the same time as providing the most significant benefit of each ESS withinside the HESS.

REFERENCES

[1] M. K. Andreev, "An Overview of Supercapacitors as New Power Sources in Hybrid Energy Storage Systems for Electric Vehicles," 2020 XI National Conference with International Participation (ELECTRONICA), 2020, pp. 1-4, doi: 10.1109/ELECTRONICA50406.2020.9305104.

[2] W. Lhomme, P. Delarue, P. Barrade, A. Bouscayrol, and A. Rufer, “Design and control of a supercapacitor storage system for traction applications,” in Conference Record of the 2005 Industry Applications Conference, Oct. 2005, pp. 2013-2020.

[3] M. Ostroverkhov and D. Trinchuk, "Increasing the Efficiency of Electric Vehicle Drives with Supercapacitors in Power Supply," 2020 IEEE 7th International Conference on Energy Smart Systems (ESS), 2020, pp. 258-261, doi: 10.1109/ESS50319.2020.9160291.

[4] Shen, J., Dusmez, S. and Khaligh, A. (2014). “Optimization of Sizing and Battery Cycle Life in Battery/Ultracapacitor Hybrid Energy Storage Systems for Electric Vehicle Applications,” IEEE Trans. Ind. Informatics, 10(4), pp. 2112–2121

[5] Araujo, R. E., Castro, R. de, Pinto, C., Melo, P. and Freitas, D. (2014). “Combined Sizing and Energy Management in EVs With Batteries and Supercapacitors,” IEEE Trans. Veh. Technol. , 63(7).

[6] [3] A. Khaligh and Z. Li, "Battery, Ultracapacitor, Fuel Cell, and Hybrid Energy Storage Systems for Electric, Hybrid Electric, Fuel Cell, and Plug-In Hybrid Electric Vehicles: State of the Art," IEEE Transaction on Vehicular Technology, vol. 59, no. 6, pp. 2806-2814, 2010.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

View publication stats

TCAD Analysis and Simulation of Double Metal Negative Capacitance FET (DM NCFET)

Yash Pathak
Department of Applied Physics
Delhi Technological University
Delhi, India
yash_2k19phdap505@dtu.ac.in

Bansi Dhar Malhotra
Department of Biotechnology
Delhi Technological University
Delhi, India
bansi.malhotra@gmail.com

Rishu Chaujar
Department of Applied Physics
Delhi Technological University
Delhi, India
chaujar.rishu@dtu.ac.in

2021 Devices for Integrated Circuit (DevIC) | 978-1-7281-9955-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/DevIC50843.2021.9455922

Abstract – **In this paper, we investigated the double metal negative capacitance field-effect-transistor (DM NCFET) for improved digital and analog performance in comparison to the conventional negative capacitance field-effect-transistor (NCFET) by using a new concept of ferroelectric layer (HfO_2FE). Visual TCAD software is used to design DM NCFET. The results have been analyzed in form of leakage current (I_{off}), transconductance (G_m), transconductance generation factor (TGF), Early Voltage (V_{EA}), intrinsic gain (A_v), unity gain cutoff frequency (f_T) and improved performance of output conductance (G_d), drain barrier induced lowering (DIBL), switching ratio I_{on} / I_{off}, subthreshold swing (SS). Thus, DM NCFET can be suitable for digital and analog circuit applications.**

Keyword – **Visual TCAD, DM NCFET, SS.**

I. Introduction

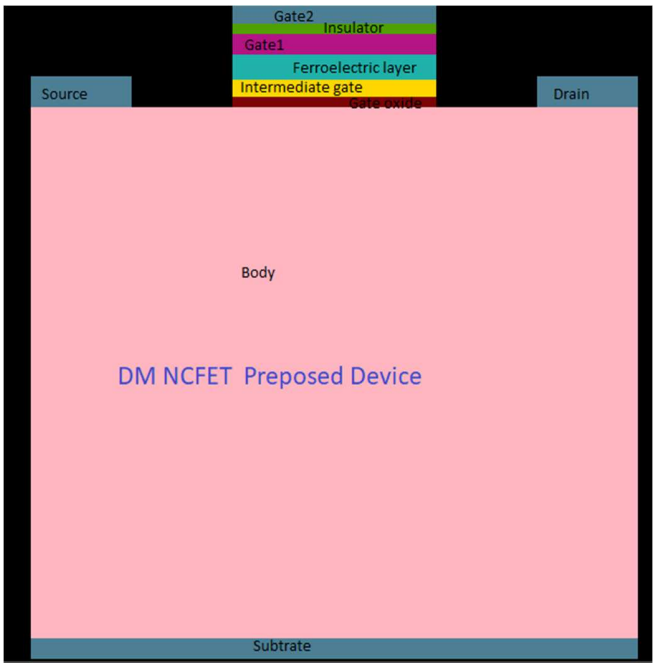
As the size of the device decreases gradually, the performance of the device gets affected and unwanted effects such as a subthreshold swing (SS), drain barrier induced lowering (DIBL) is introduced as a short channel effects (SCEs) [1] [2]. A rule of thumb that has come to dominate computing, Moore's law states that the number of transistors on a microprocessor chip will double every two years or so — which indicating that the chip performance will also perhaps double. The exponential improvement that the law describes transformed the first crude home computers of the 1970s into the sophisticated machines of the 1980s and 1990s, resulting in the rise of the high-speed Internet, smartphones and the wired-up cars, refrigerators and thermostats that are becoming prevalent today [3]. The subthreshold slope of MOSFET determines the ability to turn an on and off the device. At room temperature the ideal value is 60 mV/decade, which is determined by diffusion of carriers from the source to the channel [4]. It is generally accepted that the ongoing scaling of field effect transistors will be eventually limited by the inability to remove the heat generated in the switching process, making it very important to find ways to reduce the power dissipated per switching event. It is also clear that the power dissipation would be lowered significantly if FETs could be operated at lower voltage [5]. Landau field effect transistors promise to lower the power-dissipation of integrated circuits (ICs) by reducing the subthreshold swing (SS) below the Boltzmann limit of 60 mV/dec. The key idea is to replace the classical gate insulator with dielectrics that exhibit negative capacitance FET (NCFET) [6]. When the applied voltage is high, the charge goes down and hence, the name negative capacitance and vice -versa. It is observed from simulated result that DM NCFET have vital advantages such as decrease in leakage current, improved subthreshold swing (SS) high switching ratio, reduction in drain induced barrier lowering (DIBL) etc. with comparison to NCFET. Recently, ferroelectric FETs (Fe-FETs) have attracted enhanced attention due to their promise in both memory and switching applications [7].

II. Device Structure and Methods Simulation

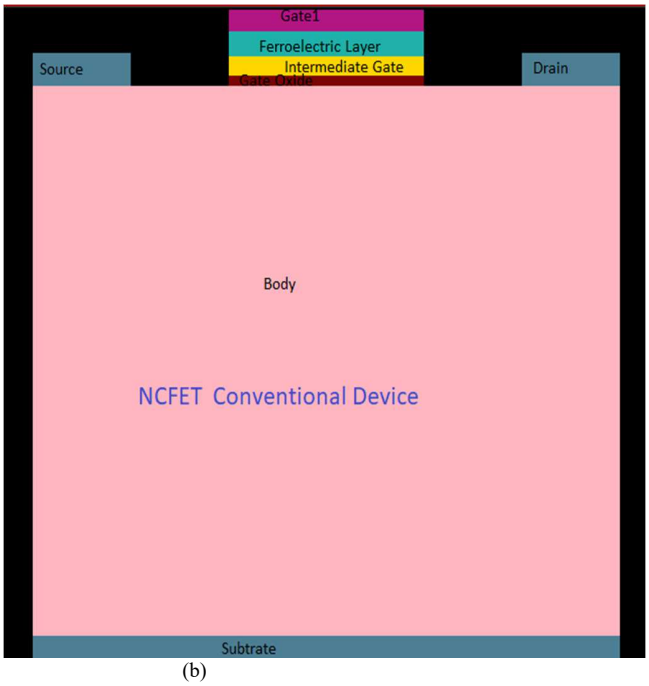
Fig. 1(a) shows the structure of a proposed double metal NCFET and Fig. 1 (b) shows the conventional device NCFET. During the entire process, the structure space and size are the same in both DM NCFET and NCFET except for the two layers. The materials used in both the devices are such that the body is made up of Si with thickness 50nm, Source /Drain are made up of Al with thickness 3nm, substrate is made up of Al with thickness 2nm, Gate oxide is made up of SiO_2 with thickness 21 nm, Intermediate gate is made up of N-poly Si with thickness 2 nm, a new ferroelectric Layer is made up of HfO_2FE with thickness 2 nm, Gate1 is made up of metal. These materials are same in both devices. The difference of proposed device from a conventional device are of two layers such that there is an insulator layer with thickness 1nm and the gate2 is made up of metal with thickness 2nm. A proposed device DM NCFET consists of metal, insulator, metal, ferroelectric, metal, insulator, oxide layer and semiconductor. All the simulations are carried out by using Visual TCAD Software. Table 2 shows the parameter, Mesh size and their value in DM NCFET. In Table 1, the doping profile and concentration of DM NCFET and NCFET device are shown. Doping of Body in DM NCFET and NCFET device is uniform with concentration 1e+16 cm^{-3}.

978-1-7281-9955-9/21/$31.00 ©2021 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:55:36 UTC from IEEE Xplore. Restrictions apply.



(a)



(b)

Fig. 1 Schematic structure of (a) Double metal NCFET Preposed device (b) NCFET conventional device.

Table 1. The doping profile used in designed DM NCFET

Name	Profile	Type	Peak conc./ cm^{-3}	Chara. /um
Body	uniform	acceptor	1e+16	-
vt	Gaussian	acceptor	1e+18	0.001
Ntype1/Ntype2	gaussian	donor	1e+20	0.0005

Table 2. Parameter and their value of DM NCFET

Parameter	Length /Thickness	Mesh size
Source /Drain	3nm	Mesh size =.01µm
Body	50nm	Mesh size =.05µm
Substrate	2nm	Mesh size =.01µm
Gate oxide	1nm	Mesh size =.0003µm
Intermediate gate	2nm	Mesh size =.001µm
Ferroelectric Layer	2nm	Mesh size =.0008µm
Gate1/Gate2	2nm	Mesh size =.01µm
Insulator	1nm	Mesh size =.001µm

III. RESULT AND DISCUSSION

Fig. 2 shows that curve of drain current, I_d (A) and Gate Voltage, V_{gs} (V) for Double Metal NCFET (DM NCFET) and NCFET at Drain Voltage V_d =0.5 V. Fig. 2 shows in log scale at Drain Voltage V_d =0.5V. In DM NCFET, the higher is the I_{on} and the lower is the leakage Current I_{off} of proposed device as compared to the conventional device.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:55:36 UTC from IEEE Xplore. Restrictions apply.

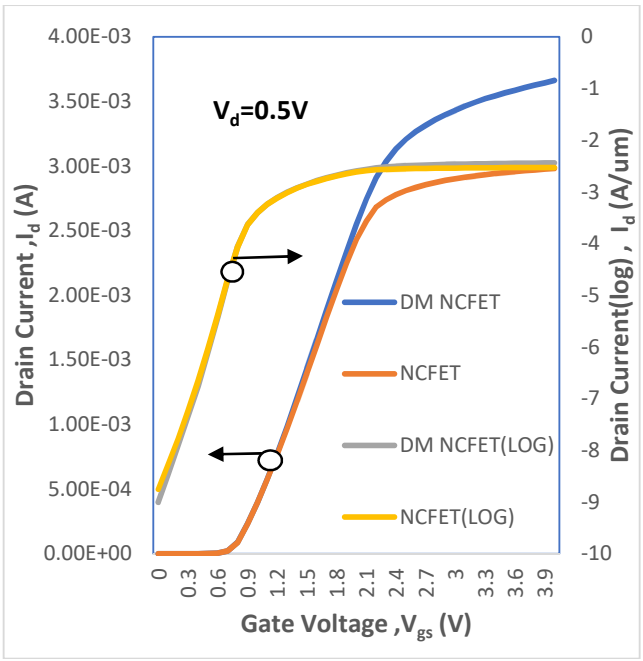


Fig. 2. Drain current and log scale vs gate voltage for DM NCFET and NCFET.

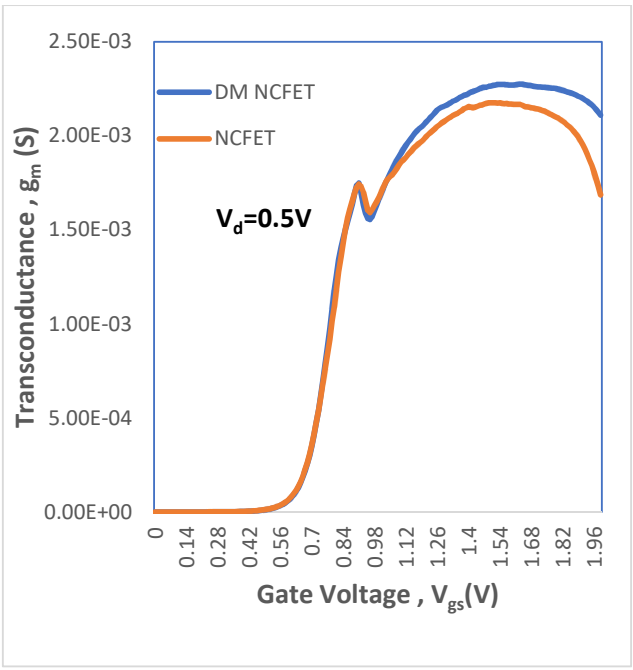


Fig. 3. Curve of Transconductance and gate voltage for DM NCFET and NCFET.

Output conductance decreases due to the shift in the lateral electric field from the gate side to the drain side, which occurs due to the increase in a fringing electric field [10]. Fig. 3. shows the curve of Transconductance and Gate voltage for Double Metal NCFET and NCFET at V_d = 0.5 V. Curve of DM NCFET is higher than NCFET revealing that the transfer of current is higher than NCFET. The transconductance increase of DM NCFET indicates increased in electronic velocity, current density and enhancing electronic mobility.

In Fig. 4 gives the curve of transconductance gain factor (TGF) and Gate voltage for Double Metal NCFET and NCFET at V_d = 0.5V. TGF is the gain generated per unit power loss. TGF perhaps will better work as low power supply. The transconductance is shown in equation 1, transconductance gain factor is shown in equation 2, output conductance shows in equation 3 current ability of device is determined by Output conductance, intrinsic gain shown in equation 4 intrinsic gain is ratio of g_m and G_d, Early voltage shows in equation 5 Early Voltage is high that means reduction in SCEs (short channel effect), Unity gain cutoff frequency is shown in equation 6. These two curves has been calculated using formulae [1] [8].

$$g_m = \frac{\partial Id}{\partial Vgs} \qquad (1)$$

$$TGF = \frac{Gm}{Id} \qquad (2)$$

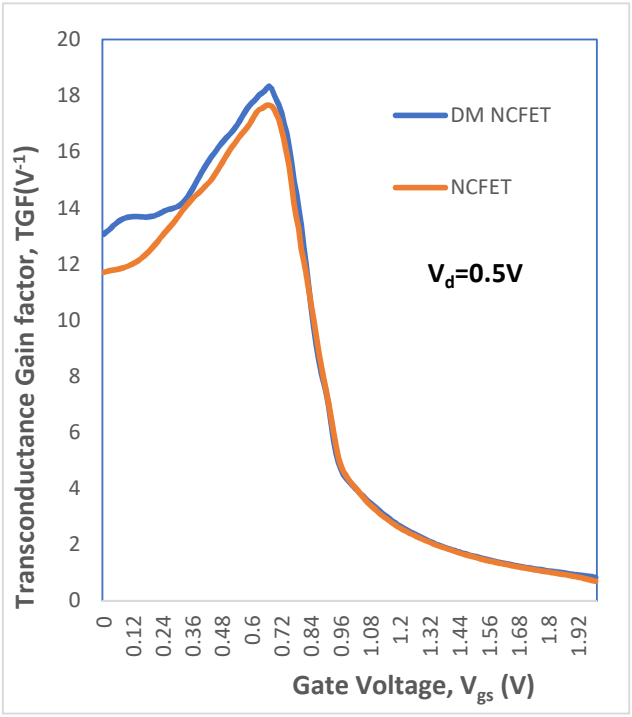


Fig. 4. Curve of TGF and gate voltage for DM NCFET and NCFET.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:55:36 UTC from IEEE Xplore. Restrictions apply.

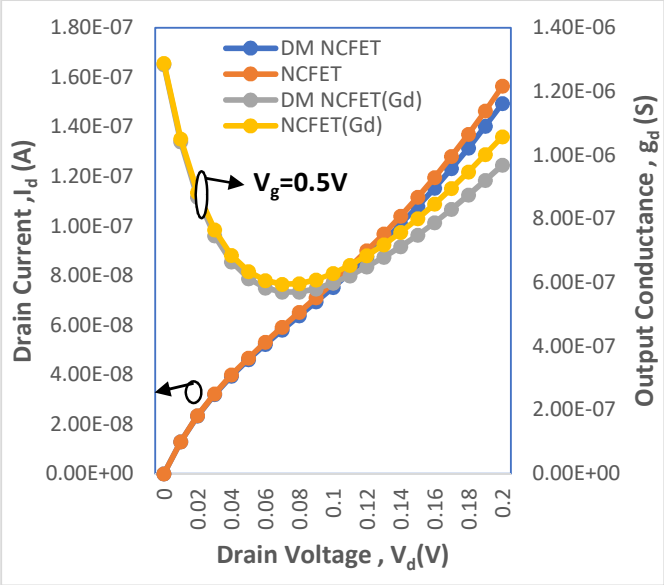


Fig. 5 Curve of Drain current and Output conductance vs Drain voltage for DM NCFET and NCFET

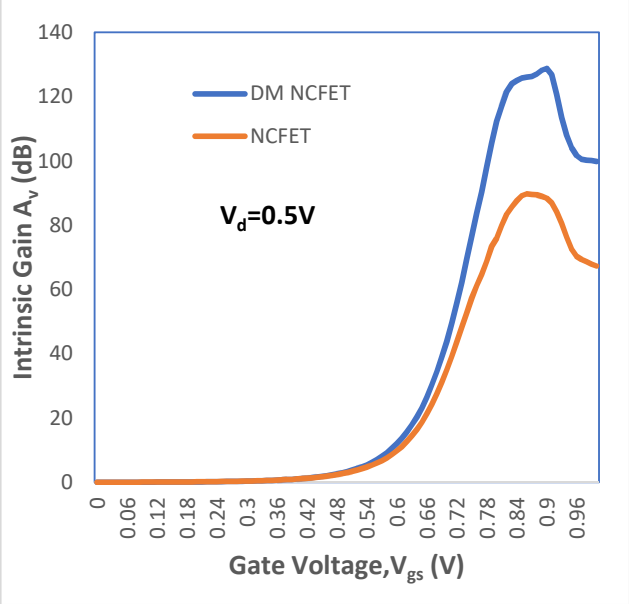


Fig. 6. Curve of Intrinsic gain and gate voltage for NCFET and DM NCFET

Fig. 5. shows the Curve of Drain current and output conductance with Drain voltage for DM NCFET and NCFET at V_{gs}=0.5V. **In** DM NCFET, the Output Conductance (g_d) is lowered than NCFET that reflects a lower Leakage current.

$$g_d = \frac{\partial Id}{\partial Vd} \qquad (3)$$

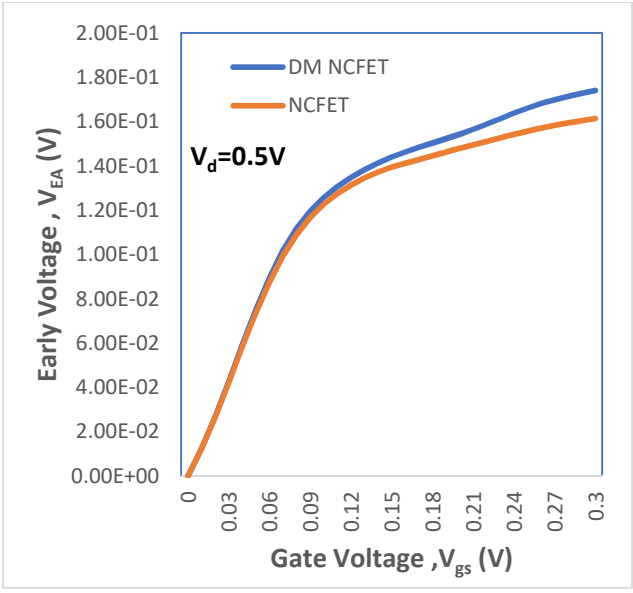


Fig. 7. Curve of Early voltage and gate voltage for DM NCFET and NCFET

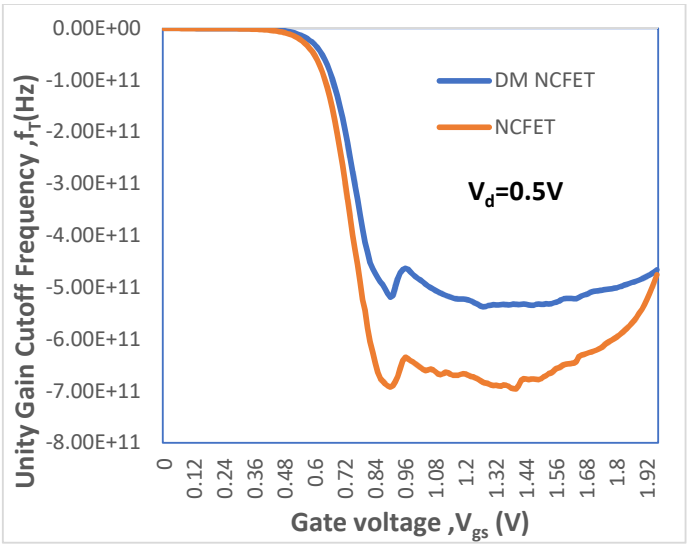


Fig. 8. Curve of Unity gain factor and gate voltage for DM NCFET and NCFET

Fig. 6. shows the curve of intrinsic gain and gate voltage for DM NCFET and NCFET at V_d =0.5V. Intrinsic gain is RF (Radio frequency) performance of device. Intrinsic gain of DM NCFET is higher than NCFET which means a higher conductance gain of device. All the three parameter have been calculated by using the following equations [9].

$$A_v = \frac{g_m}{g_d} \qquad (4)$$

$$V_{EA} = \frac{I_d}{g_d} \qquad (5)$$

$$f_T = \frac{g_m}{2\pi(C_{gs}+C_{ds})} \qquad (6)$$

here, C_{gs} shows capacitance gate to source
C_{gd} shows capacitance gate to drain

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:55:36 UTC from IEEE Xplore. Restrictions apply.

Table 3. Value of Digital and Analog Performance for DM NCFET and NCFET.

S.No.	Device Name	Vth lin(V)	Vth sat(V)	I_{on} (A/um)	I_{off} (A/um)	I_{on}/I_{off}	SS	DIBL
1.	DM-NCFET	0.7105	0.6183	0.002528	$9.28*10^{-10}$	$2.585*10^{6}$	165.954mV/Decade	0.2304
2.	NCFET	0.7108	0.61227	0.002415	$1.76*10^{-9}$	$1.37*10^{6}$	174.869 mV/Decade	0.246

The Fig. 7. variation of early voltage and gate voltage for DM NCFET and NCFET at V_d=0.5V that reflects the RF performance of device. Early voltage is inversely proportional to output conductance and is higher than NCFET that reveals reduction in SCEs. The Fig. 8. exhibits the curve of unity gain cutoff Frequency and Gate voltage for NCFET and DM NCFET at V_d=0.5V, It is inversely proportional to capacitance and also a parameter of RF performance. The curve of unity gain cutoff frequency of DM NCFET is higher than NCFET that reflects lower capacitance. In the Table 3. Represent that reduced subthreshold swing (SS) – control of the channel improved I_{on} /I_{off} ratio and - improved leakage current and decrease in drain induced barrier lowering (DIBL). All the simulation is obtained from Visual TCAD Software. I_{off} current (leakage current) of DM NCFET is lower than that of NCFET.

IV. Conclusion

In this work, the effect of ferroelectric thin layer (HfO_2FE) in negative capacitance Field Effect Transistor (NCFET) is investigated by using Visual TCAD Software. From the simulated result, it has been shown that Double Metal NCFET is improved more than NCFET. The proposed device (DM NCFET) has decreased leakage current (I_{off}) by 47 % than that of the conventional device (NCFET). This device has decreased in sub threshold swing (SS) by 5 % than that of the conventional device (NCFET). the proposed device has also decreased in Drain induced barrier lowering (DIBL) by 6% than that of NCFET. These result clearly reveal that DM NCFET has better and improved the digital and analog parameter than that of NCFET.

References

[1] B. Kumar, A. Kumar and R. Chaujar, "The Effect of Gate Stack and High-κ Spacer on Device," in *IEEE VLSI Device, Circuit and System Conference (VLSI-DCS)*, Kolkata, India, 2020, pp. 159-163.

[2] A. Chabbra, A. Kumar and R. Chaujar, "GaAs Junctionless FinFET using Si3N4 Spacer for High performance Analog Application," in *International Conference On Advances in Communication and Computing Technology (ICACCT)*, 2018, pp. 483-486.

[3] M. M. Waldrop, "The chips are down for Moore's law," *Nature,* vol. 530, pp. 144-147, 2016.

[4] K. E. Moselund, D. Bouvet, V. Pott, C. Meinen, M. Kayal and A. M. Lonescu, "Punch-through impact ionization MOSFET (PIMOS): From device principleto applications," *Solid-State Electronics,* vol. 52(9), pp. 1336-1344, 2008.

[5] S. Salahuddin and S. Datta, "Use of Negative Capacitance to ProvideVoltage Amplification for Low PowerNanoscale Devices," *NANO LETTERS,* vol. 8, pp. 405-410, 2007.

[6] K. Karda, A. Jain, C. Mouli and M. A. Alam, "An anti-ferroelectric gated Landau transistor to achieve sub-60 mV/dec switching at low voltage and high speed," in *Applied Physics Letters 106(16) ,163501*, 2015.

[7] G. A. Salvatore, L. Lattanzio, D. Bouvet, I. Stolichnov, N. Sstter and A. M. Lonescu, "Ferroelectric transistors with improved characteristics at high temperature," in *Applied Physics Letters 97, 053503*, 2010.

[8] S. K. Mohapatra, K. P. Pradhan, P. K. Sahu and M. R. Kumar, "The performance measure of GS-DG MOSFET: An impact of metal gate work function," in *Adv. Nat. Sci. Nanosci. Nanotechnol., vol. 5, no. 2,*, 2014.

[9] B. Kumar and R. Chaujar, "Analog and RF Performance Evaluation of Junctionless Accumulation Mode(JAM) Gate Stack Gate All Around (GS-GAA) FinFET," in *Silicon*, 2021.

[10] A. Nandi, A. K. Saxena and S. Dasgupta, "Impact of dual-k spacer on analog performance of underlap FinFET," *Microelectronics J.,* vol. 43, no. 11, pp. 883-887, 2012.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on July 07,2021 at 05:55:36 UTC from IEEE Xplore. Restrictions apply.