SCHOLARLY PUBLICATIONS

A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT

@DTU

(108th Edition)

DECEMBER 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 **One Hundred Eight** 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 **December, 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. A Bridgeless Isolated Zeta Converter for Battery Charging Aided with Solar Photovoltaic System, ***7.Chanchal Chaudhary***, ***3.Saurabh Mishra*** and ***3.Uma Nangia***, Electrical, DTU

2. A Review of Web Infodemic Analysis and Detection Trends across Multi-modalities using Deep Neural Networks, ***7.Chahat Raj*** and ***3.Priyanka Meel***, IT, DTU

3. Analysis of MISO Converter with Intelligent Control for Solar Water Pumping, ***7.Shivangi Yadav***, ***3.Saurabh Mishra*** and ***3.Garima***, Electrical, DTU

4. ANALYSIS ON FLOATING SOLAR POWER PLANT, ***3.Dr. Jp Kesari***, ***8.Abhay Pratap Singh Shishodia*** and ***8.Eshaan Gupta***, Mechanical and Applied Physics, DTU

5. ANGULAR ESTIMATES FOR CERTAIN ANALYTIC FUNCTIONS, ***8.POOJA YADAV*** AND ***3.S. SIVAPRASAD KUMAR***, APPLIED PHYSICS, DTU

6. ANN & PSO based Approach for Solar Energy Forecasting; A Step Towards Sustainable Power Generation, ***6.Md Tabish Ansari*** and ***3.M. Rizwan,*** Electrical, DTU

7. Barriers to corporate social responsibility implementation in the medium size manufacturing sector: an interpretive structure modelling approach, ***6.Gaurav Dawar*** and ***3.Seema Singh***, Humanities, DTU

8. Career Recommender System Using Decision Trees, Massoud Massoudi, Siyamoy Ghory and ***7.Mahboob Massoudi***, CSE, DTU

9. Clickbait in YouTube: Prevention, Detection and Analysis of the Bait using Ensemble Learning, ***8.Peya Mowar***, ***8.Mini Jain***, ***8.Ruchika Goel*** and ***3.Dinesh Kumar Vishwakarma***, IT, DTU

10. Comparative Study of Deep Learning Models for Devanagari OCR, ***Shilpa Kaur Manocha*** and ***3.Piyush Tewari***, Electronics, DTU

11. Comparative Study Of Impact Of COVID-19 Lockdown On AQI Parameters Across Urban India, ***8.1.Himanshu Nandanwar*** and ***3.Anamika Chauhan***, IT, DTU

12. CREB1K292 and HINFPK330 as Putative Common Therapeutic Targets in Alzheimer's and Parkinson's Disease, ***8.Rohan Gupta*** and ***3.Pravir Kumar***, Biotechnology, DTU

13. Dear technology, you are not welcome: exploring ungratefulness towards technology, ***3.Naval Garg*** and Sarika Kumari, DSM, DTU

14. De-CAPTCHA: A novel DFS based approach to solve CAPTCHA schemes, Aditya Atri, ***6.1.Ankita Bansal***, Manju Khari and S. Vimal, CSE, DTU

15. Devanagari Handwritten Character Recognition using CNN as Feature Extractor, ***6.Shilpa Kaur Manocha*** and ***3.Piyush Tewari***, Electronics, DTU

16. EASE Em: Hybrid Recommendation System for Item Cold Start Problem, ***8.S. Aditya***, ***8.Mihir Rajora*** and ***3.Indu Singh***, CSE, DTU

17. Effect of Shear Walls on Tall Buildings with Different Corner Configuration Subjected to Wind Loads, ***8.Saransh Mahajan***, ***8.Vikramaditya Yadav***, ***8.Rahul Raj*** and ***3.Ritu Raj***, Civil, DTU

18. Exploring Linkages Between Workplace Spirituality and Occupational Stress Among Indian Police Personnel, Stuti Jalan and ***3.Naval Garg***, DSM, DTU

19. FPGA Based Comparative Analysis Of Single Carrier PWM Schemes, ***6.Yatharth Gupta*** and ***3.Madhusudan Singh***, Electrical, DTU

20. Frequency analysis of multi-source single area power system using WCA algorithm, ***7.Kirtiman Godara*** and ***3.Narendra Kumar***, Electrical, DTU

21. Gene expression based inference of drug resistance in cancer, Smriti Chawla, Anja Rockstroh, Melanie Lehman, Ellca Rather, Atishay Jain, Anuneet Anand, ***8.Apoorva Gupta***, Namrata Bhattacharya, Sarita Poonia, Priyadarshini Rai1, Nirjhar Das, Angshul Majumdar, Jayadeva, Gaurav Ahuja, Brett G. Hollier, Colleen C. Nelson and Debarka Sengupta, Biotechnology, DTU

22. Grey Wolf Optimizer-Based Array Reconfiguration to Enhance Power Production from Solar Photovoltaic Plants under Different Scenarios, ***6.Astitva Kumar***, ***3.Mohammad Rizwan***, ***3.Uma Nangia*** and Muhannad Alaraj, Electrical, DTU

23. Handling class imbalance problem in software maintainability prediction: an empirical investigation, ***3.Ruchika MALHOTRA*** and ***3.Kusum LATA***, CSE and USME, DTU

24. Insights into Structural Modifications of Valproic Acid and Their Pharmacological Profile, ***6.Manish Kumar Mishra***, Samiksha Kukal, Priyanka Rani Paul, ***6.Shivangi Bora***, Anju Singh, Shrikant Kukreti, Luciano Saso, Karthikeyan Muthusamy, ***3.Yasha Hasija*** and Ritushree Kukreti, Biotechnology, DTU

25. Investigation of charge transport mechanism in hydrothermally synthesized reduced graphene oxide (rGO) incorporated zinc oxide (ZnO) nanocomposite films, ***Nikita Jain***, Savita Sharma and ***3.Nitin K. Puri***, Applied Physics, DTU

26. Link Prediction in Social Networks using Histogram Based Gradient Boosting Regression Tree, ***7.Himanshu Tiwari*** and ***3.Sanjay Kumar***, CSE, DTU

27. Metasurface Superstrate Beam Steering Antenna with AMC for 5G/WiMAX/WLAN Applications, ***6.Akhilesh Verma***, Ravi Kumar Arya and ***3.Srinivasa Nallanthighal Raghava***, DTU

28. Modeling and Control of PANTHERA Self-Reconfigurable Pavement Sweeping Robot under Actuator Constraints, ***3.Madan Mohan Rayguru***, M. R. Elara, A. A. Hayat, B. Ramalingam and Spandan Roy, Electrical, DTU

29. Neuro-fuzzy network incorporating multiple lexicons for social sentiment analysis, ***6.Srishti Vashishtha*** and ***3.Seba Susan***, DTU

30. Performance analysis of inclined microjet impingement heat sink with porous medium, ***8.Harshit Jain***, ***8.Kanishk Yadav***, ***8.Kapeesh Kumar*** and ***3.M. Zunaid***, Mechanical, DTU

31. Performance Evaluation of Three-Phase Unified Power Quality Conditioner controlled using SRF, ***8.1.Rahul Singh*** and ***3.Alka Singh***, Electrical, DTU

32. Plasma-based nanoarchitectonics for vertically aligned dual-metal carbon nanotube field-effect transistor (VA-DMCNFET) device: effect of plasma parameters on transistor properties, ***6.Mansha Kansal*** & ***3.Suresh C. Sharma***, Applied Physics, DTU

33. Rethinking Neural Networks with Benford's Law, Surya Kant Sahu, ***8.Abhinav Java*** and Arshad Shaikh, CSE, DTU

34. Retinal Vessel Detection Using Residual Y-net, ***7.Anindita Roy*** and ***3.Kapil Sharma***, IT, DTU

35. Sentiment Analysis on Twitter using Ordinal Regression, ***8.Moin Ahmed***, ***8.Mohit Goel***, ***8.Raju Kumar*** and ***3.Dr. Aruna Bhat***, CSE, DTU

36. Significant Improvement in Magnetic and Magnetoelectric Characteristics of (0.95-x)Ba0.9Ca0.1TiO3-0.05Na0.5Bi0.5TiO3- xCoFe2O4 ceramics, Aditya Jain, Ajay Kumar and ***6.1.Neha Gupta***, Applied Physics, DTU

37. SSR Stability Analysis and Modeling for DFIG Connected to Series Compensated Transmission Line, ***6.Neevatika Verma*** and ***3.Narendra Kumar***, Electrical, DTU

38. Statistical Characterization of the Angle of Arrival in an Underwater Channel due to scattering from the sea surface modeled as Sum of Sinusoids, Manishika Rawat, ***7.1.Brejesh Lall*** and Seshan Srirangarajan, Electronics, DTU

39. Stumpy-Paced Green Power Generation System, Ashok Kumar Singh, P.Kumar, R.Kishore, ***6.Phool Singh***, ***6.Neeraj Kant***, ***3.Pushpendra Singh***, Gopal Nandan, A. Jaiswal and R. Prasad, Mechanical, DTU

40. Temperature-Dependent Electric Field-Induced Optical Transitions of 2D Molybdenum Disulfide (MoS2) Thin Films: Temperature- Dependent Electroabsorption and Absorption, ***3.Mohan Singh Mehata***, ***6.Mrityunjay Kumar Singh***, Kamlesh Awasthi, ***7.Prateek Sharma***, Shailesh Rana and Nobuhiro Ohta, Applied Physics, DTU

41. Theoretical Study of Highly Efficient CH3NH3SnI3 based Perovskite Solar Cell with CuInS2Quantum Dot, Gagandeep, ***3.Mukhtiyar Singh***, Ramesh Kumar ,Vinamrita Singh and Sunita Srivastava, Physics, 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*

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)
Pune, India. Oct 29-30, 2021

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON) | 978-1-6654-2503-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/SMARTGENCON51891.2021.9645803

A Bridgeless Isolated Zeta Converter for Battery Charging Aided with Solar Photovoltaic System

Chanchal Chaudhary
Electrical Engineering Department
Delhi Technological University
Shahbad Daulatpur, Main Bawana Road, Delhi NCR-110042
Delhi, India
chanchalchaudhary94@gmail.com

Saurabh Mishra
Electrical Engineering Department
Delhi Technological University
Shahbad Daulatpur, Main Bawana Road, Delhi NCR-110042
Delhi, India
saurabhmishra@dtu.ac.in

Uma Nangia
Electrical Engineering Department
Delhi Technological University
Shahbad Daulatpur, Main Bawana Road, Delhi NCR-110042
Delhi, India
uma_nangia@rediffmail.com

***Abstract*— The idea of assisting solar PV in onboard battery charging system improves reliability of system consequently less burden on grid and improvement in per mile cost of vehicle is achieved and so an assisted charging system including bridgeless isolated Zeta converter as power factor correction (PFC) unit is introduced here. The converter operation is ensured in discontinuous conduction mode (DCM) for wide operating range whereas CUK converter having continuous output and input current features is used in solar PV is utilized efficiently for obtaining maximum power from solar PV. Further when the grid outage occurs, additional benefits include continuous battery charging, supply voltage control for wide variation. Also, system is observed at various irradiations in solar PV and results are presented here. Supply current is improved by using PFC and absence of diode bridge rectifier in bridgeless isolated Zeta converter makes it ripple free circuit by reducing switching stresses thereby decreasing Total harmonic distortion (THD) value to standard acceptable range i.e., below 5% showing improved and efficient charging system. Cascaded PI control is used to ensure the closed loop operation in PFC converter while solar PV uses the MPPT algorithm as P&O to extract the maximum of power from the PV panel and ensuring the continuity of battery charging in smooth way. Moreover, the PFC converter circuit uses CC-CV charging. The proposed system is validated using MATLAB Simulink.**

Keywords*— *Bridgeless isolated Zeta, onboard charging, Power factor correction, CUK converter, Light electric vehicle (LEV).

I. INTRODUCTION

Light electric vehicles (LEV) cover a wide portion in total count of the electric vehicles and becomes unavoidable while exploring electric vehicles and its charging operation. In general examples of LEV are E-cycle, E-bikes, E-rickshaw, E-scooter etc. In charging system of LEV's problems arises due to the distortion in supply (Harmonics due to usage of power electronic switching converters) causing low power factor, high distortions at input supply current and high values of total harmonic distortions (THD) as per [1]. Consequently, the efficiency of the charger deteriorates and the purpose of economic benefit from LEV's remains unfulfilled. The offboard charging stations were already established widely and now onboard charging is opening new area for researchers as making the LEV's more cost efficient and reliable can directly benefit the people related to this industry. That's why this is a keen research focused area in EV. However, some problems like very low efficiency of chargers due to supply side distortions can make it difficult to design such systems where continuous charging of battery can be done in effective manner.

Though many techniques have been used nowadays to improve charger properties, one of them is using power factor correction (PFC) unit. PFC unit can be any DC-DC converter like Buck, Boost or Buck-Boost, Cuk, Zeta, Luo, SEPIC etc. as mentioned in [1] and [2]. Out of these converters Buck-Boost converter is preferred over Buck and Boost converters due to advantage of the wide variation in duty cycle of former over other two. Although with the introduction of some new converters, usage of buck-boost is reduced. As per [3] CUK, Zeta, Luo are efficient converters as PFC unit than buck-boost converters because of better current profiles in these converters but above all nowadays some techniques like bridgeless converters [4], Interleaved converters [5], isolated converters [6] or their combinations are excellent choices in the PFC phenomenon of battery charging as they provide advantages like in interleaved converter advantage of self-cancellation of ripple within the circuit is there. Moreover, bridgeless converters are also useful in many applications like in BLDC drives, monitoring operation in hospitals etc. These pros in new converters are giving good results where THD is reduced to higher extent as compared to conventional topologies, power factor is improved to a great extent, harmonics at input mains current reduces largely as mentioned in [7]. Despite of these methods integrated systems are not explored much yet so it is a new and effective method to enhance system effectiveness by some experimentation with solar PV as it is an easily available source of energy and renewable also. Although efficiency of generation is low but still it can add on to the charging.

The efficient system is designed to fulfill the charging needs of battery from supply as well as from renewable source (solar). This is useful at the time when power is not present for charging the battery and also at the time when power is there it can still help in enhancing the overall charging of the system. One of this kind of system is discussed in [8]. This paper reflects the system of bridgeless isolated Zeta converter [9] and solar fed CUK converter [8] charging the Li-ion battery. Converter being used in DCM mode to make certain and sound operation of PFC converter. The bridgeless property ensures the reduction in conduction losses at the input which were previously caused due to the diode bridge rectifier while isolation property in Zeta converter is helpful in eliminating the ripples to larger

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:00:31 UTC from IEEE Xplore. Restrictions apply.

extent making the system efficient thereby cutting down the THD of input supply current to less than 5%. This reduction in THD results in the wave shaping of input current and is helpful in making the power factor closer to unity. PID controlling technique is discussed in [10], in this paper PI technique is used to control the voltage and current of the battery. In solar fed system having CUK converter, P&O is used for the tracking of maximum power due to its certain advantages over previously used MPPT techniques i.e., it has low burden of computational losses as well as the tracking speed is very high. It is here employed to track the maximum power from the solar PV at different levels of irradiations and the results are shown in further discussed sections. The whole system has been validated in Simulink for observing the performance of battery charging by both bridgeless isolated Zeta converter as well as by solar fed CUK converter.

II. Configuration and Working Operation

The configuration here deals with an onboard charger with bridgeless isolated Zeta converter as PFC unit integrated with solar PV for battery charging. The solar PV integrated system can charge the battery in the absence of AC supply as well as it can boost the charging of battery in presence of supply and PFC circuit. In solar PV circuit, CUK converter steps down the DC voltage generated at the PV input to the required level with the help of variation in duty cycle of switch S_V . Fig.1. shows the proposed configuration that how integrated system works to fulfill the reliable system needs.

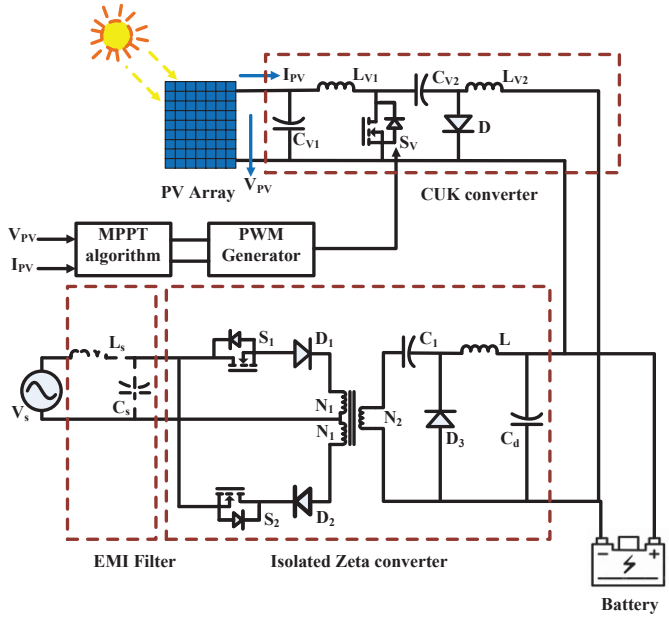


Fig.1. Proposed Configuration

Working Operation

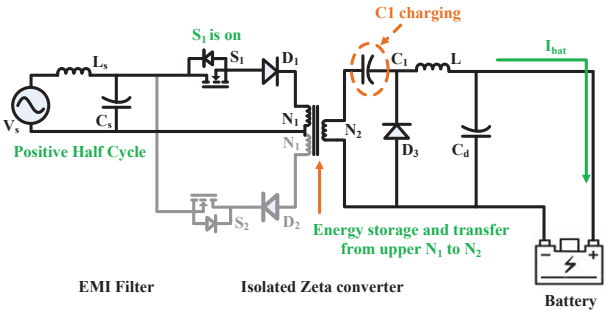
A. PFC converter-bridgeless isolated Zeta

The PFC converter works in three modes broadly on the basis of how switching of S_1 and S_2 works. Isolation polarity decides the on and off working of the components at output side of bridgeless isolated Zeta converter. The designed converter is made to work in DCM mode.

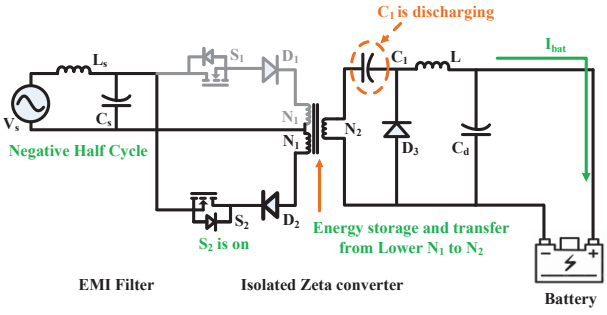
Mode 1- Fig.2. (a) and Fig.2. (b) represents mode 1 during positive and negative half cycles respectively. In positive half cycle switch S_1 becomes on and consequently upper half of the transformer is in working mode i.e., energy is stored in magnetizing inductance causing capacitor C_1 to be in charging mode whereas diode D_3 is in non-conducting state. Battery is charged via C_1 and L as illustrated in Fig.2. (a). In case of Fig.2. (b) during negative cycle switch S_2 is on and upper half is non-working. Diode D_3 is in forward biased and in working condition. Charging voltage is supplied by the output capacitor in this mode.

Mode2- After mode 1 both the switches are made to turn off in mode 2 as shown in Fig.2. (c). In this charging of the battery is ensured by the output inductor as both the switches are in off condition and magnetizing inductance is also getting discharged.

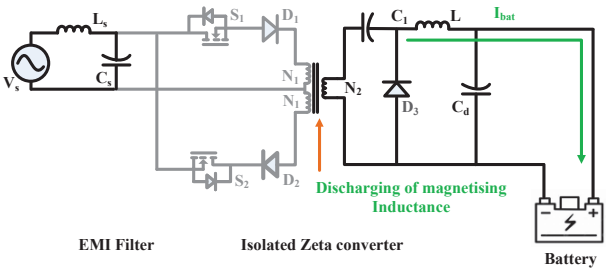
Mode3- The magnetizing inductance is completely discharged in mode 3 and isolation is created from primary side of the transformer as can be seen in Fig.2. (d) and the charging of battery is now completely the task of output inductor.



(a)

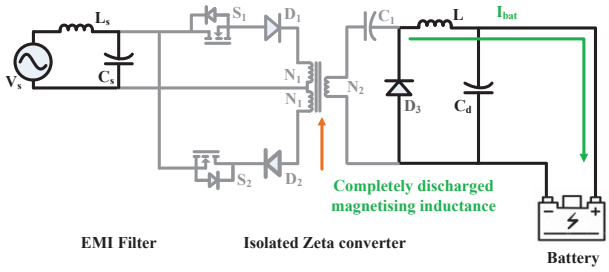


(b)



(c)

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:00:31 UTC from IEEE Xplore. Restrictions apply.



(d)

Fig.2. (a) Mode 1 in positive half cycle (b) Mode 1 in negative half cycle (c) Mode 2 (d) Mode 3

III. Design Considerations

A battery of 48V is getting charged with the help of PFC fed circuit as well as from solar fed circuit having CUK converter in it. The bridgeless isolated zeta converter is supplied by input supply as defined by:

$V_S(t) = V_m \sin \omega t$; where ω is $2\pi f$ and frequency f is given by 50 Hz. Output equation for charging is given by:

$$V_o = \left(\frac{N_2}{N_1}\right)\left(\frac{D}{1-D}\right)V_{in} \tag{1}$$

where $\left(\frac{N_2}{N_1}\right)$ is the transformer turns ratio considered as 1:2 here and V_{in} is the average voltage given as:

$$V_{in} = \frac{2V_S}{\pi} \tag{2}$$

The critical magnetizing inductance value can be calculated as per the given equation (3) and the chosen inductance value should be very lower than the critical value.

$$L_m = \frac{\left(V_o^2/P\right)(1-D)^2}{2Df_s\left(N_2/N_1\right)^2} \tag{3}$$

where f_s is PFC converter's switching frequency. The capacitor C_1 present at the middle of the circuit and is found with the help of equation (4).

$$C_1 = \frac{V_o D}{\Delta V_{C_1} f_s \left(V_o^2/P\right)} \tag{4}$$

The inductor L at output has specifications given by:

$$L = \frac{V_o(1-D)}{f_s \Delta I_L} \tag{5}$$

where, ΔI_L is the ripple in inductor current and it is regarded as 10% of inductor current. Output inductor value is high as it fulfills the energy requirement of load during the entire switching process. The DC link capacitor C_d at the output of PFC converter is calculated with the help of equation (6).

$$C_d = \frac{I_o}{2\omega \Delta V_d} \tag{6}$$

where ΔV_d is the ripple in DC link's voltage which is chosen as 2% of the DC link voltage of capacitor.

Filter capacitance C_s and L_s are given by the equations (7) and (8) respectively. The calculated value is the maximum value of filter capacitance and chosen value should be less than calculated one.

$$C_s = \frac{P\sqrt{2}/V_S}{\omega V_m}\tan\theta \tag{8}$$

Where θ is the angle of displacement in supply voltage and current at mains.

$$L_s = \frac{1}{4\pi^2 f_c^2 C_s} \tag{9}$$

Where f_c is filter's cut off frequency which is chosen as $1/10^{th}$ of the converter frequency.

All the calculated values of Bridgeless isolated Zeta are given in Appendix as Table. I.

The design of CUK converter used in Solar PV circuit for getting the desired voltage for battery charging are given as follows: Inductors L_{v1} and L_{v2} are calculated by the formulas as per equations (10) and (11):

$$L_{v1} = \frac{(1-D)}{2Df}\left(V_o^2/P\right) \tag{10}$$

$$L_{v2} = \frac{(1-D)}{2f}\left(V_o^2/P\right) \tag{11}$$

The capacitors C_{v1} and C_{v2} are given by the equations (12) and (13).

$$C_{v1} = \frac{DV_o}{\Delta V_{c_{v1}} f\left(V_o^2/P\right)} \tag{12}$$

$$C_{v2} = \frac{(1-D)V_o}{8\Delta V_{c_{v2}} f^2 L_{v2}} \tag{13}$$

All the calculated parameters of CUK converter have been mentioned in Appendix as Table. II.

IV. Control Strategy

Isolated Zeta converter has two switches S_1 and S_2 which are controlled using cascaded PI loop as shown in Fig.3., in which battery voltage is then compared to a voltage that has been taken as a reference voltage and voltage controller is fed with error voltage. The output signal of voltage PI controller is compared with battery current and is fed to the current PI controller which in turn generates pulses with the help of PWM generator. For PI tuning method opted here is Zeigler Nicholas. Before finding the values of K_p and K_I it is important to obtain transfer function of the system which has been achieved with the help of steady state modelling and using state space method. After finding out transfer function, Routh Hurwitz method is used to find out the critical value of K_{cr} afterwards equations (13), (14) and (15) can be used to find out the values of K_p and K_I.

$$T_{cr} = \frac{2\pi}{\omega_{cr}} \tag{13}$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:00:31 UTC from IEEE Xplore. Restrictions apply.

$$K_I = \frac{1.2K_P}{T_{cr}} \quad (14)$$

$$K_P = 0.45K_{cr} \quad (15)$$

In solar PV system, the switch S_v is controlled with the assistance of maximum power point tracking (MPPT) algorithm which is P&O technique here. The flow chart in Fig.4. shows the working of this technique in proper manner. Maximum power from PV array is extracted by using CUK converter as shown in Fig.1. and variation in duty cycle is used to achieve the maximum point of power in P-V curve at the concerned irradiation.

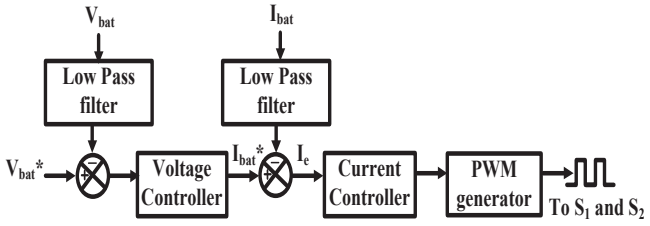


Fig.3. Control for Isolated Zeta converter

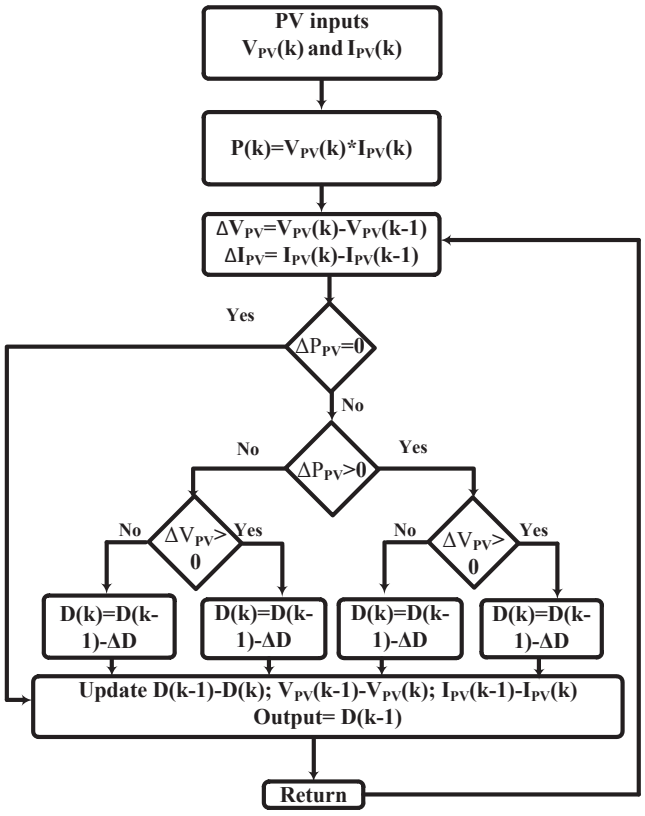


Fig.4. P&O Algorithm

V. RESULTS

The results for PFC Zeta converter are as shown in Fig.4. It can be clearly observed that supply voltage and supply current are in phase showing the power factor closer to unity. Constant battery voltage is showing the charging of battery constantly. Negative battery current is a sign of charging of battery and hence li-ion battery of 48V,100Ah is getting charged with a constant current of 10A. Increasing SOC% is another parameter showing charging of battery as in Fig.4. THD of input supply current is found to be 1.30% which represents reduction in ripple to a large extent in the circuit. Overall battery is getting efficiently charged with the help of bridgeless PFC Zeta converter. In Fig.6. charging via solar fed CUK converter has been illustrated at solar irradiation of 1000W/m^2 while Fig.7. shows the variation of charging the battery by PFC assisted with solar at an irradiation of 1000W/m^2. At various irradiations charging of battery is observed and assisted system has been found effective in improving the charging properties of the circuit.

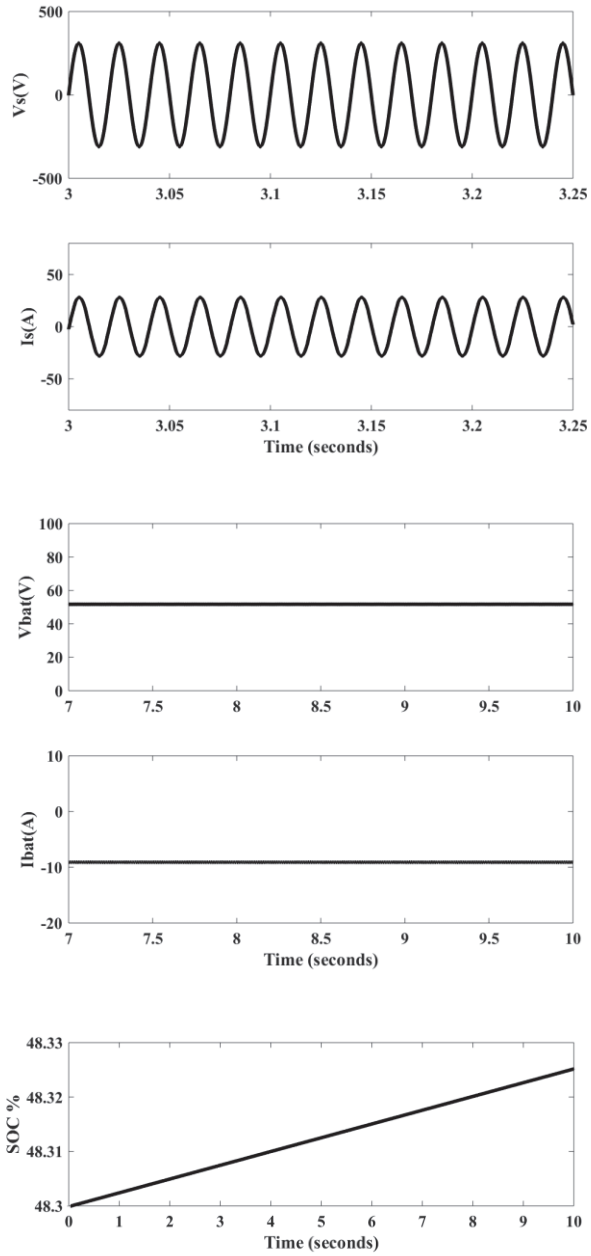


Fig.4. Charging characteristics of PFC converter: Input Voltage, Input Current, battery voltage, battery current, SOC%

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:00:31 UTC from IEEE Xplore. Restrictions apply.

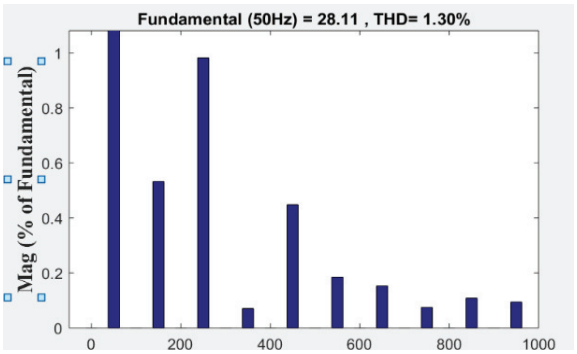


Fig.5. THD of input Supply

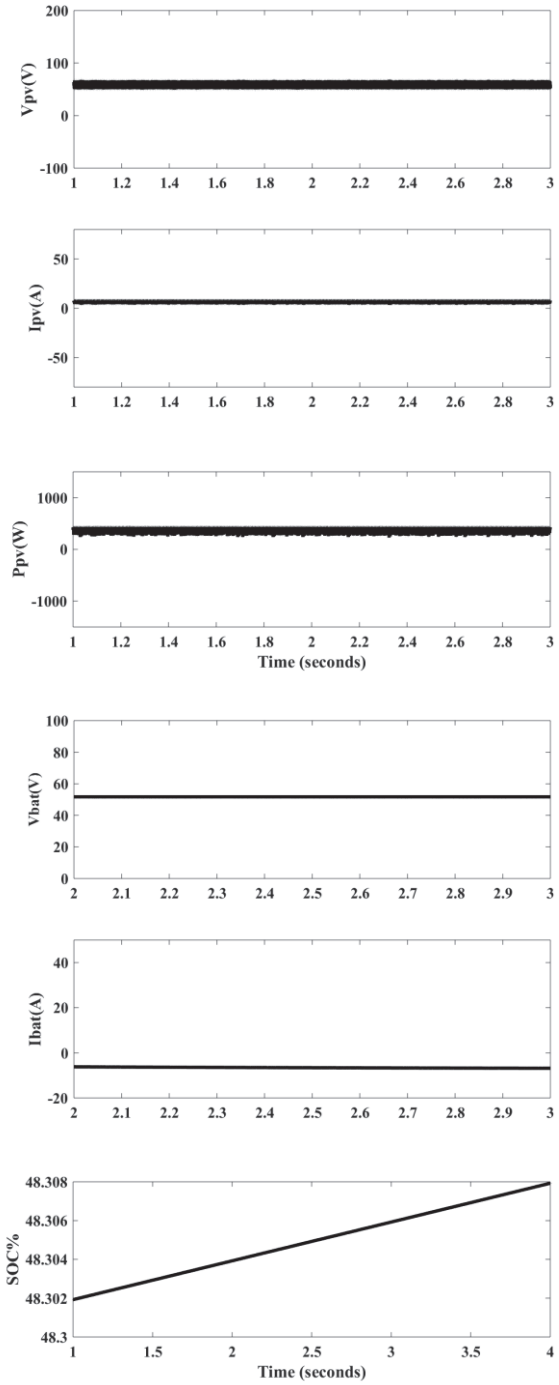


Fig.6. Charging characteristics of solar fed CUK converter charging at solar irradiation 1000W/m^2: Vpv (V), Ipv(A), Ppv(W), Vbat(V), Ibat(A), SOC%

PFC is assisted with solar for first five seconds and other five seconds only PFC circuit is working. Sunrays falling on the panel and the amount of panel area exposed to the sun decides the temperature and irradiation of the system. Here in this system results are observed at solar irradiation (1000W/m^2) and a temperature of 25 degrees Celsius. Solar fed CUK converter showing charging of battery at an irradiation of 1000W/m^2 is shown in Fig.7.

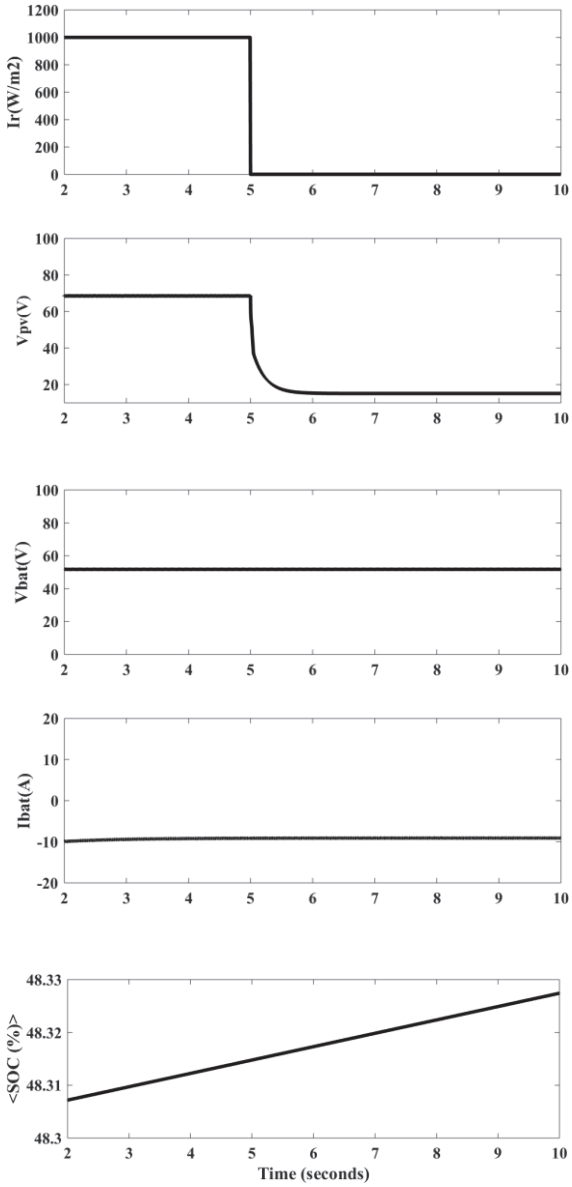


Fig.7. Charging characteristics of solar fed CUK converter charging with PFC assisted with solar irradiation 1000W/m^2 for 0-5s Ir(W/m2), Vpv (V), Vbat(V), Ibat(A), SOC%

VI. Conclusion

An assisted solar PV charging has been presented in this paper which is an efficient and cost-effective method for the charging of battery in LEV's. The effectiveness and reliability of system has been improved by utilizing the grid power as well as solar power to enhance the battery charging. THD of the system has been reduced to 1.30% while using PFC bridgeless isolated Zeta converter and in addition to this power factor is closer to unity while battery is getting charged with this prototype.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:00:31 UTC from IEEE Xplore. Restrictions apply.

ACKNOWLEDGMENT

The author is thankful for all the help and support to Electrical Engineering Department of Delhi Technological University (DTU) and also acknowledges SMARGEN 2021 for providing me opportunity to present my findings.

APPENDIX

TABLE I. BRIDGELESS ISOLATED ZETA CONVERTER

S.No.	*Parameters*	*Values chosen*
1.	Supply voltage V_s	220V
2.	Average output voltage V_{in}	198V
3.	Switching frequency f_s	20kHz
4.	Filter capacitance C_s	330nF
5.	Filter Inductance L_s	3mH
6.	Output Capacitor C_1	330nF
7.	Output Inductor L	0.8mH
8.	DC link capacitor C_d	10mH
9.	Magnetizing Inductance L_m	50µH

TABLE II. CUK CONVERTER PARAMETERS

S.No.	*Parameters*	*Values chosen*
1.	Input Inductor L_{v1}	2.3mH
2.	Output Inductor L_{v2}	2.5mH
3.	Input Capacitor C_{v1}	100µF
4.	Output Capacitor C_{v2}	5µF

REFERENCES

[1] J. Gupta, R. Kushwaha and B. Singh, "Improved Power Quality Transformerless Single-Stage Bridgeless Converter Based Charger for Light Electric Vehicles," in *IEEE Transactions on Power Electronics*, vol. 36, no. 7, pp. 7716-7724, July 2021.

[2] S. S. Indalkar and A. Sabnis, "Comparison of AC-DC Converter Topologies Used for Battery Charging in Electric Vehicle," 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), pp. 1624-1627, July 2019.

[3] A. Jha and B. Singh, "Portable Battery Charger for Electric Vehicles," *International Conference on Sustainable Energy and Future Electric Transportation (SEFET),* pp. 1-6, Jan 2021

[4] R. Kushwaha, B. Singh, "Design and Development of Modified BL Luo Converter for PQ Improvement in EV Charger," *in IEEE Transactions on Industry Applications*, vol. 56, no. 4, pp. 3976-3984, July-Aug 2020.

[5] R. Kushwaha and B. Singh, "Interleaved Landsman Converter Fed EV Battery Charger with Power Factor Correction*," IEEE 8th Power India International Conference (PIICON*), Dec 2018.

[6] K. Mishra, M. K. Pathak and S. Das, "Isolated converter topologies for power factor correction — A comparison*," International Conference on Energy, Automation and Signal*, Bhubaneswar, pp. 1-6, 2011

[7] R. Kushwaha and B. Singh, "A Modified Bridgeless Cuk Converter based EV Charger with Improved Power Quality," *IEEE Transportation Electrification Conference and Expo (ITEC), Detroit, MI, USA*, pp. 1-6, June 2019.

[8] J. Gupta, B. Singh and R. Kushwaha, "A Rooftop Solar PV Assisted On-Board Enhanced Power Quality Charging System for E-Rickshaw," *International Conference on Power, Instrumentation, Control and Computing (PICC),*Dec 2020.

[9] B. Singh and R. Kushwaha, "An EV battery charger with power factor corrected bridgeless zeta converter topology*," 7th India International Conference on Power Electronics (IICPE)*, pp. 1-6, Nov 2016.

[10] N. Tiwari and A. N. Tiwari, "Performance Analysis of Unidirectional and Bidirectional Buck-Boost Converter Using PID Controller," 2nd *International Conference on Electronics, Materials Engineering & Nano-Technology*, pp. 1-6, may 2018.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:00:31 UTC from IEEE Xplore. Restrictions apply.

A Review of Web Infodemic Analysis and Detection Trends across Multi-modalities using Deep Neural Networks

Chahat Raj[1], Priyanka Meel[2]
Department of Information Technology
Delhi Technological University-110042, New Delhi, India
[1]chahatraj58@gmail.com, [2]priyankameel86@gmail.com

Abstract: Fake news and misinformation is a matter of concern for people around the globe. Users of the internet and social media sites encounter content with false information much frequently. Fake news detection is one of the most analyzed and prominent areas of research. These detection techniques apply popular machine learning and deep learning algorithms. Previous work in this domain covers fake news detection vastly among text circulating online. Platforms that have extensively been observed and analyzed include news websites and Twitter. Facebook, Reddit, WhatsApp, YouTube, and other social applications are gradually gaining attention in this emerging field. Researchers are analyzing online data based on multiple modalities composed of text, image, video, speech, and other contributing factors. The combination of various modalities has resulted in efficient fake news detection. At present, there is an abundance of surveys consolidating textual fake news detection algorithms. This review primarily deals with multi-modal fake news detection techniques that include images, videos, and their combinations with text. We provide a comprehensive literature survey of eighty articles presenting state-of-the-art detection techniques, thereby identifying research gaps and building a pathway for researchers to further advance this domain.

Keywords: Fake news detection, Visual data, Multi-modal, Deep learning, Neural Networks

1. Introduction

Since the advent of online social platforms, the need to establish baselines for information-assessment across all information flow channels has demanded researchers' attention. It has been dubious how wisely people using these platforms actively communicate and digest the information circulating on the internet. With the exponential emergence of social media as a globally used platform in the recent decade, we have encountered massive escalation in fake news dispersion.

Any act of deliberate, miscreant, or unverified inclusion of information creates fake news. Substantial instances include widespread misinformation since January 2020 claiming World War 3 [1], its probability, countries involved, and tentative dates. The arrival of the COVID-19 pandemic has led to a multi-fold rise in disseminating fake news globally. This escalated spread of misinformation amidst the pandemic has been termed as 'Infodemic.' Statements such as "Turmeric powder and black pepper cure coronavirus [2]", "Cocaine treats COVID [3]" etc. have appeared as post contents on every social site in the forms of text messages, images, and videos. Such fake news can be found listed in various official websites debunking the claims [4]. Another fake message states a WHO-issued four-step protocol to prevent COVID-19 circulated at large on most online social networking platforms, especially on WhatsApp [5]. Conspiracy theories and tweets conveyed that the stated countries created the novel coronavirus as a weapon to ignite bio-war [6]. Subsequently, fake news about the pandemic became as contagious as the virus itself. A huge source of fake news was the 2016 US presidential elections that have also focused on many fake news researchers. Other such events and examples of fake news are Pizzagate [7], Indian and Brazilian elections [8], Hurricane Sandy [9], spying technology in Rs. 2000 Indian currency notes [10], Citizenship Amendment Act 2019 [11] & Article 370 (Kashmir, India) [12], etc., which have had huge adversarial impacts on the lives of people.

Figure 1: Visual fake news example of a shark in Houston (Source: USAToday)

Social media interactions have contributed a lot towards inventories of big data. Users accord it in text, image, video, audio, emoticons, reactions, etc. Text, images, and videos compose a major actuating portion of the information on the internet and hence blot the grey areas of fake

news. Since fake news and its critical nature came into the picture, several scientific developments have been made in textual fake news detection. Fake news detection technologies using NLP, text classification, vector-space models, rhetoric structure theory, opinion mining, sentiment analysis, graph theory, deep neural networks, and others have been created, reviewed, and summarized by fellow researchers [13-16]. Probing of image fake news is comparatively lower, and of videos, it is negligible. Most of the information users encounter on the internet is accompanied by visual representations, either using an image, video, or other modalities. Visual data is quickly gazed upon and leaves a lasting impact. It is the breeding ground of tampering, manipulations, and forgeries. With technology emanating, editing applications and techniques, images and videos are being meddled to mislead information consumers. Moreover, online social media allows users to add to their data, be it real or fake. Recent examples include an image of a shark on a Houston freeway that went viral during Hurricane Harvey (Figure 1). The image was shared and retweeted at large, creating havoc amongst people. Misinformation has been so outrageous that even US President Donald Trump could not escape from it and went on to retweet a fake video that stated anti-malaria drugs could cure coronavirus, which was later brought down when proven fake [17]. In 2020, during the coronavirus pandemic, fake news claiming Kim Jong Un dead, absconding, or assassinated spread widely [18]. Fake videos of his funeral were shared on social media (Figure 2).

Figure 2: Outdated image of King Jong Un's Death (Source: nypost.com)

In this survey, we broadly categorize data modalities into text, images, and video that are spreading fake news on the internet. We present a literature review in fake news detection techniques that

cover these data modalities individually or in combination as hybrid or multi-modal data. This review intends to cover detection mechanisms for data modalities that promote most of the spreading fake news, i.e., text, images, and videos. This survey aims to highlight detection mechanisms that can detect misinformation based on any of these data types. We enlist the challenges and limitations of past research and enlighten ways to future work.

1.1 Motivation

As we observe trends among fake news detection, a huge amount of published work and resources are available for text-based detection. Considerable amounts of frameworks and detection mechanisms using textual features have been designed since the problem domain emerged. Machine learning and deep learning algorithms have been applied largely to provide solutions due to their extreme popularity in this domain. These include sentiment analysis, text mining, stance classification, similarity analysis, etc. Texts are analyzed based on their sentence structures, words, punctuations, tone, grammar, and pragmatics. Textual fake news analysis is the domain well explored and worked upon by a large number of researchers. Various other detection techniques have emerged, like using psycholinguistic features, deception modeling, fake news spreading prediction, graph-based propagation detection, and reputation scores.

We present this review in a modality-focused manner. Fake news spreads in the form of textual or visual data. We aim to elaborate on how mechanisms have been developed to deal with each type of fake information that spreads textually through images or videos. The fake news domain has been expressively presented by Sharma and Sharma [19], Figueira and Oliveira [20], Torabi and Taboada [21], Zhang and Ghorbani [22], Tandoc et al. [23], and Shu et al. [24]. Summaries of noteworthy works have been provided in reviews by Mosinzova et al. [25] and Rubin et al. [26]. In contrast, Rajendran et al. [27] brought into the light the importance of Deep Neural Networks for stance classification.

Moving towards fake visual news, tasks performed have been comparatively lower in number. The domain started gaining attention in 2013 and rose considerably from 2017 to 2020. Image fake news detection started gaining importance when image-accompanied news and posts started appearing at large on online platforms. The rise in visual data made ways for fake news to seep in and thus encouraged researchers to explore this domain. Cao et al. [28] have highlighted the role of exploring visual data while detecting fake news. Fake news detection has shown notable performance improvements when a visual analysis is combined with text. Most of the visual

classification techniques applied neural networks that provided quick and efficient results. Fake images can be classified as tampered images where some of the other manipulations are made in the image or as misleading images where the context of news content and image do not comply. Some fake news is also accompanied by older images from other events, i.e., combining recent news with outdated images. Another category of fake images that recently emerged is computer-generated images commonly created by Generative Adversarial Networks (GANs). All these types of images contribute to fake news. Any FND framework has not been able to detect fake news that revers to all these types collectively. Individual modules have been created for different tasks. Some architectures can classify tampered images well, while others can spot misleading images that do not match the context. Some frameworks classify images based on statistical features of textual and visual data. Parikh and Atrey [29] provided a survey for multimedia FND. There is a lack of a wholesome framework that could efficiently detect fake news under all modalities. As we talk of visual data, fake news videos have taken up at large over social media. These have a great impact on the minds of viewers and bring adversarial social and political effects. Frameworks for video fake news detection are very few. Multiple features are needed to be studied for video analysis, being a complex task containing spatial and temporal features, speech, and movements. Few researchers have applied it using inconsistencies between speech and lip movements; few have attended analyzing facial expressions following similar trends in real and fake videos, while few utilize the image information at every frame. Other multimedia news in audio, podcasts, and broadcasts are yet less infiltrated with fake news. With an acutely low amount of work performed, video fake news detection is an emerging problem domain that researchers need to pay heed to.

1.2 Existing Challenges in Multimodal Fake News Detection

The domain of fake news detection gained popularity very quickly. Large amounts of unverified and uncredible posts have been misleading people. Using linguistic features for credibility assessment of content is a popular and widely-used method. Here, we list the existing challenges to detect fake news spreading through all types of data.

1. **Visual Fake News Detection:** The fake news menace is rising. It began with spreading through text and has now started gripping users through all forms of multimedia. Visual data probable to fake news exploitation can be categorized into images or videos. There are various existing approaches to detect text-based fake news. However, visuals play a great role in

impacting viewers' minds and therefore are being infiltrated with fake news in the current generation. An image or video can be easily modified using media editing applications. Various manipulations in visual data go unrecognized through the viewers' eyes. It is very difficult for humans to observe minor changes in modified images and videos to classify them as fake or real. Automated tools are required which can identify minute variations created made to fabricate or manipulate visual data. This poses a great challenge for researchers in designing visual-based fake news detectors.

2. **Auditory Fake News Detection:** Auditory fake news is a type of fake news that has been in existence but is going unnoticed. Many social applications allow users to share recorded audios. These audios files are vulnerable to spreading fake content, propaganda, unverified information, and more. This type of multimedia has not been put into use for credibility assessment. There are no fake news detection mechanisms that incorporate audio as a sole modality or in combination with other data modalities. This issue needs to be addressed to prevent the contamination of audios with false content and serve its early detection.

3. **Detection of Embedded Fake Content:** When one type or modality of data is fixed or embedded within another type of data, it is embedded content. A new type of social media posts is spreading widely known as a 'meme.' It is an image or a video, mostly with text embedded on it. Various forms of media like text, image video, gifs, or hyperlinks are embedded into other forms. It is a complex task to analyze media that is embedded in some other data type. There is an upsurge of fake content in the form of embedded media or memes. Efficient detection mechanisms are required to fight such misleading content.

4. **Multimodal Datasets:** To build fake news detection mechanisms, most machine learning and deep learning tasks require large amounts of data. There is a lack of real-world multimodal datasets. Text-based fake news datasets are more in number than visual or multimodal datasets. Lack of proper datasets limits the extent of research. There is a need to collect real-world fake news data that consists of various types of information like text, image, video, and meta-data.

5. **Holistic Detection Mechanism:** The research community has encountered many techniques that can robustly detect fake news. These techniques use linguistic features, visual features, sentiment scores, social context, network/propagation-based features, meta-data, and hybrid features. There is no such mechanism at present that extracts all these details from a given fake content and predicts its integrity based on all the contributing factors. Different researchers have highlighted the importance of all of these techniques individually or in hybrid combinations. It is worthwhile to consider all these features for building a holistic fake news detector.

6. **Real-Time Verification:** Provided information-spread ease through the internet and online social platforms, fake news is being generated and spread at every instant. Fake news can be about anything and anyone. It spreads continuously as we interact on the internet. Existing detection tools either require users to self-validate a piece of news by fact-checking on their website/application or classify news late after it has been spread and affected various aspects of life. The world needs a system that analyses content in real-time and declares it as fake or real based on its decision.

7. **Lack of Literature:** There is a lack of significant literature in the domain of multi-modal fake news detection. Although many authors have presented textual detection mechanisms, work done in the multimodal domain is minimal and not complete. In this paper, we endeavour to cover all the past research performed using multiple data modalities other than only text-based techniques. We highlight works that have used visual content, alone or along with textual content, for detecting fake news.

1.3 Review methodology and data analysis

We provide an enriched and systematic literature review of all the relevant and available articles in the multi-modal and visual fake news detection domain. The articles have been extracted from reputed digital libraries that include Google Scholar, IEEE Xplore, Elsevier, Springer, ACM, and Wiley.

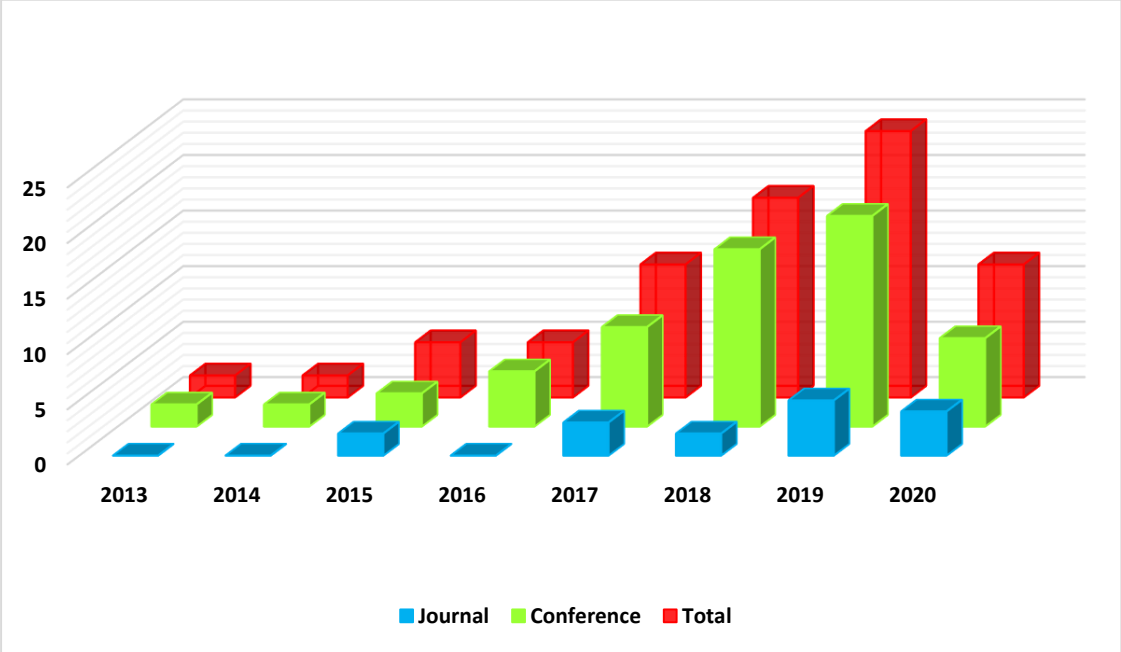


Figure 3: Year-wise trend of published work

We studied over 200 articles on the topic from the libraries mentioned above and shortlisted 70 relevant articles suitable for this survey that display state-of-the-art performance. We have covered all important articles and literature published in various journals, conferences, book chapters, and student thesis reports. Figure 3 provides an overview of the trend of works published yearly between 2013 and 2020. All the related research has been performed, and relevant articles have been published in this span. Some of the brilliant works done have been published in reputed journals or conferences. From the figure, we can observe that focuses moved towards visual and multi-modal fake news detection largely after 2016. The number of published works has been rising since then. It is easily recognizable that tasks involving visual data in the fake news domain has been rising for the past five years and is grabbing researchers' attention.

The search words applied for querying digital databases include multi-modal fake news detection, image fake news detection, fake news detection, multi-modal fake news datasets, and their synonyms. Figure 4, presents the article-wise percentage distribution of algorithms and techniques utilized for multi-modal fake news detection.

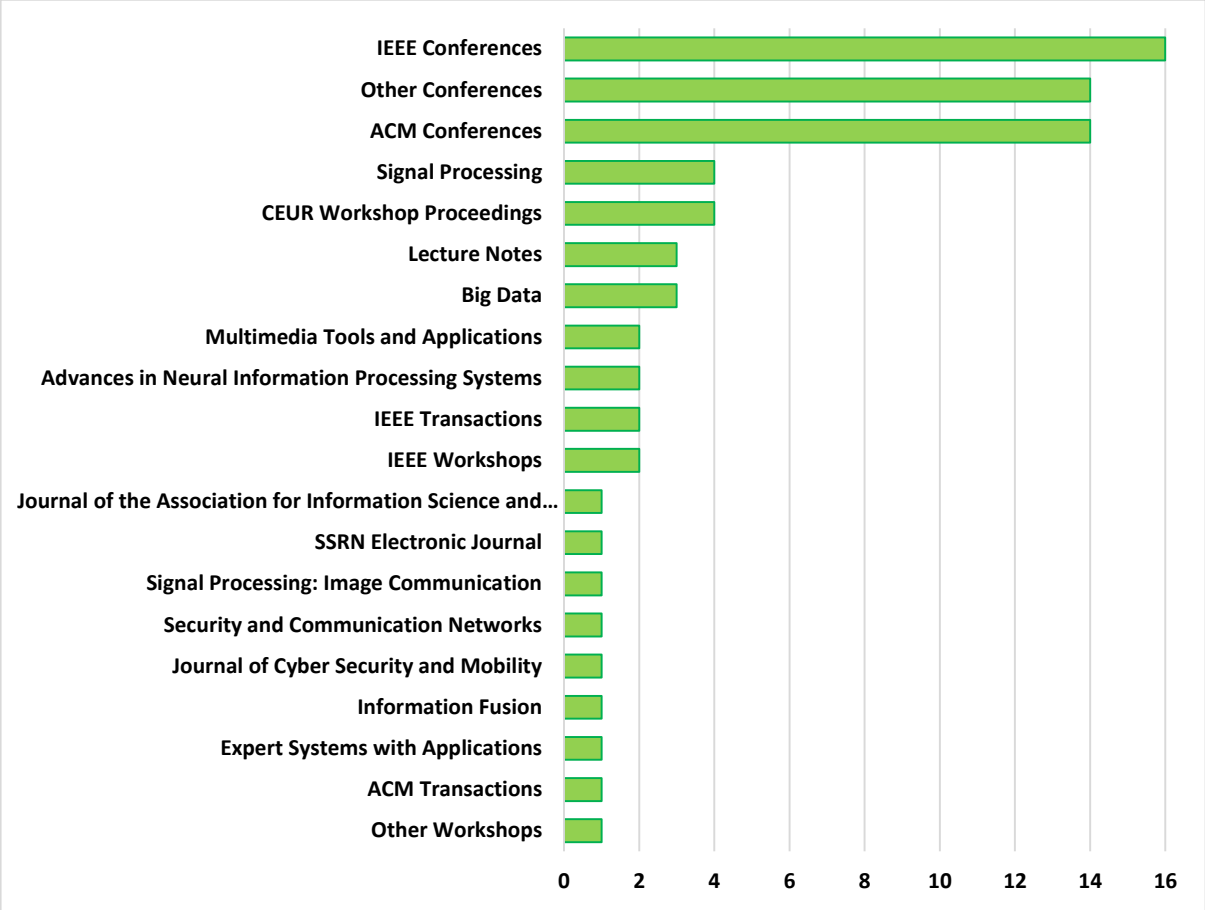


Figure 4: Number of articles published in journals, conferences, and lecture notes

1.4 Contribution and Organization of this paper

To the best of our knowledge, this is the first data modality-based review in fake news detection. The contributions of this study are as follows:

- Analyzing and identifying the techniques utilized in multi-modal fake news detection tasks.
- Comparing these techniques based on their applications, advantages, and disadvantages.
- Providing a comprehensive review of remarkable work done in the domain, discussing popular techniques, datasets used, and results obtained.
- Providing a detailed summary of multi-modal usable datasets for fake news detection.
- Comparing the efficiencies of available literature and their work in terms of evaluation parameters utilized.
- Identifying the research gaps in multimodal fake news detection methods and enlisting potential research directions.

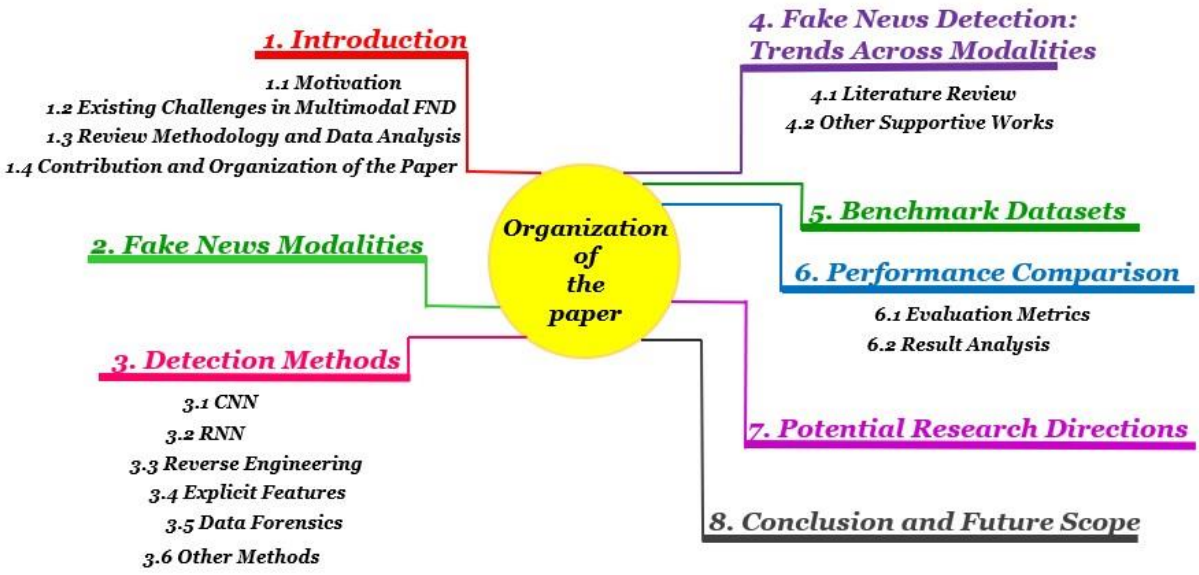


Figure 5: Organization of the paper

The organization of this paper is depicted in figure 5. Section 1 discusses the current Infodemic situation and introduces readers to the problem domain. We describe the motivation to conduct this review and the existing challenges in multimodal fake news detection. The review methodology used and analysis of articles is described next. In section 2, we describe the types of modalities in which fake news propagates, which forms the basis of this review. Section 3 discusseses several detection methods used so far across textual and visual modalities and list their advantages, disadvantages, and applications. Section 4 provides a detailed review of the literature and important architectures built for fake news detection tasks. Section 5 serves the readers with rich tabular information of benchmark multi-modal datasets for FND tasks. Section 6 explains the performance evaluation metrics various researchers have used in their works, their distribution, and performance analysis of noteworthy architectures. In section 7, we discuss potential future research directions. Section 8 concludes this review by summarizing our work and imparting motivation and future research directions to readers.

2. Fake News Modalities

Fake news is defined as any piece of false information that misleads people. It can be deliberate, fabricated, or simply unintentional. The intent of spreading false news could be maliciously intentional, political, for gaining monetary benefits, popularity, or simply for fun. While referring to data modalities, fake news spreads through text, images, videos, audios,

hyperlinks, embedded content, and hybrids. Because of less or no work in the remaining modalities, research is limited to textual and visual modalities. Therefore, we consider these modalities for review, which have been explored by researchers for fake news detection.

Text: This is the most popular mode of communication on the internet. People interact through textual matter on social media platforms, websites, blogs, e-mails, personal messaging, and more. Most of the false information spreads through text on the internet. Fake news is found propagating on social media posts, articles, and online messaging services. Text is the simplest and most used way for an internet user to convey his concerns. Being a largely used modality for communication, it also accounts for a large amount of fake news. Figure 6 shows an example of fake textual news. The screengrab is taken from Twitter. In the tweet, the user falsely attributes a claim to WION News, which states that China is hiding the real numbers of death amidst the coronavirus pandemic. The post says that SO2 concentration around Wuhan, China, has grown due to the burning of a large number of dead bodies. The claim is false and has been debunked by various fact-checking websites.

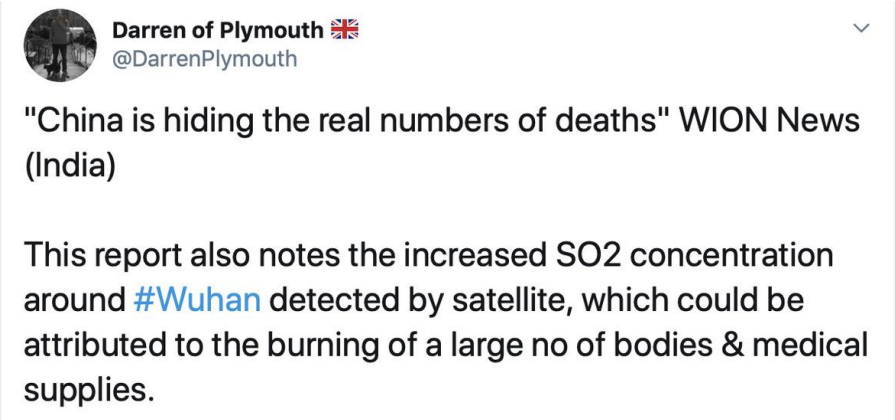


Figure 6: Textual fake news example (Source: Twitter)

Image: An image is a visual representation of something. Images are highly vivacious and impactful for depicting anything. They leave lasting impressions on the minds of viewers, whereas words can be forgotten shortly. They have suddenly gained popularity with the increase in the feasibility of sharing them. Images go through certain manipulations to carry a false message. The use of photo editing tools supports these manipulations. Some examples of editing techniques are cropping, splicing, copy-move, retouching, or blurring. Any image can be manipulated to convey a false message, which contributes to fake news. Often, they are not manipulated but accompanied

by false text. Many of the times, irrelevant or out-of-context image is placed with fake text. All of these types of images emanate false conceptions accord with fake news. In figure 7, a Facebook post shows a girl rescuing a koala bear from Australia's bushfires. Originally, the picture is a digitally created artwork used out-of-context to match the bushfire situation.

Figure 7: Example of fake news image (Source: Facebook)

Video: Sharing of videos online became immensely easy with the introduction of YouTube. The platform allows the feasible sharing of information through video content. Social media platforms allow various techniques for communication through videos. This can be done in the form of regular posts, stories, ads, or even comments. This viability of video interaction gives room for sharing of fake content through videos. Videos are a powerful and impactful tool. They are capable of successfully manipulating people through their content. Therefore, it raises a serious concern to authenticate video content and decide whether a video is credible or not. Figure 8 shows a screengrab from Twitter that shows a video claiming Vladimir Putin’s daughter was getting the first shot of coronavirus vaccine. The girl is, in fact, a volunteer and not the Russian President’s daughter.

Figure 8: Example of fake news video (Source: Twitter)

Other Modalities: Numerous data types have not been analyzed for fake news yet due to limitations of exposure and datasets. These include embedded data types, audios, and hyperlinks (clickbait). Embedded media is that where one type of data is merged or superimposed onto another. For example, textual matter on images or videos, embedding audio in images, altering audio in videos, etc. Detection of fraudulent content in such a data type is complex and challenging. There is a lack of past research in this area. Figure 9 shows a meme with a text embedded on it that says that the North Korean leader Kim Jong Un faked his death to expose traitors. Many such false statements and claims circled the internet. Various fact-checking sites have debunked these claims.

Figure 9: Fake example of a text-embedded image (Source: Facebook)

3. Detection Mechanisms

Various fake news detection algorithms utilized by researchers are explained below. Figure 10 depicts the percentage distribution of used algorithms and methods in the reviewed articles. These include Deep Neural Networks, Image/Video Forensics, explicit features, and other methods. Textual detection mechanisms have employed several machine learning and deep learning classifiers. They have been successful enough in this domain. For visual analysis, the most popular tools are Deep Neural Networks. Many researchers have combined these with the use of explicit features available on the web. These can be statistical features, user-based features, post-based features, propagation features, or more. Another popular method is the use of reverse image search on search engines. This method is useful in identifying the integrity of a particular image. Tools are available that use this mechanism to allow the verification of fake news. Several other methods of image and video manipulation have started being merged with fake news detection. This has bridged the gap between methodologies and brought the problem domain and possible solutions in a common range. The applications, advantages, and disadvantages of the following methods have been summarized in table 1.

3.1 CNN

First introduced in the 1980s, CNNs have come a long way in the domain of computer vision. They have been applied to Natural Language Processing [30], image classification [31], video classification [32], object recognition [33], time-series forecasting [34], anomaly detection [35], speech analysis [36], handwriting recognition [37] and the likes. For image classification, CNNs require training over large image datasets. Their learning process occurred to be substantially faster than previous methods known, an underlying feature that brought CNNs into the picture. They are efficient in analyzing latent features present inside an image or a video. A rich survey of the latest Convolutional Neural Networks has been provided in Khan et al. [38]. As compared to images, little work has been done in video classification using CNNs as videos are more complex to process owing to their temporal dimension. Most works utilize CNNs to classify videos by extracting images at every video frame [39]. Another method treats spatial and temporal domain separately and classify them using two convolutional neural networks and fusing them after that. Many researchers have also applied CNNs for text classification using one-dimensional convolutional networks.

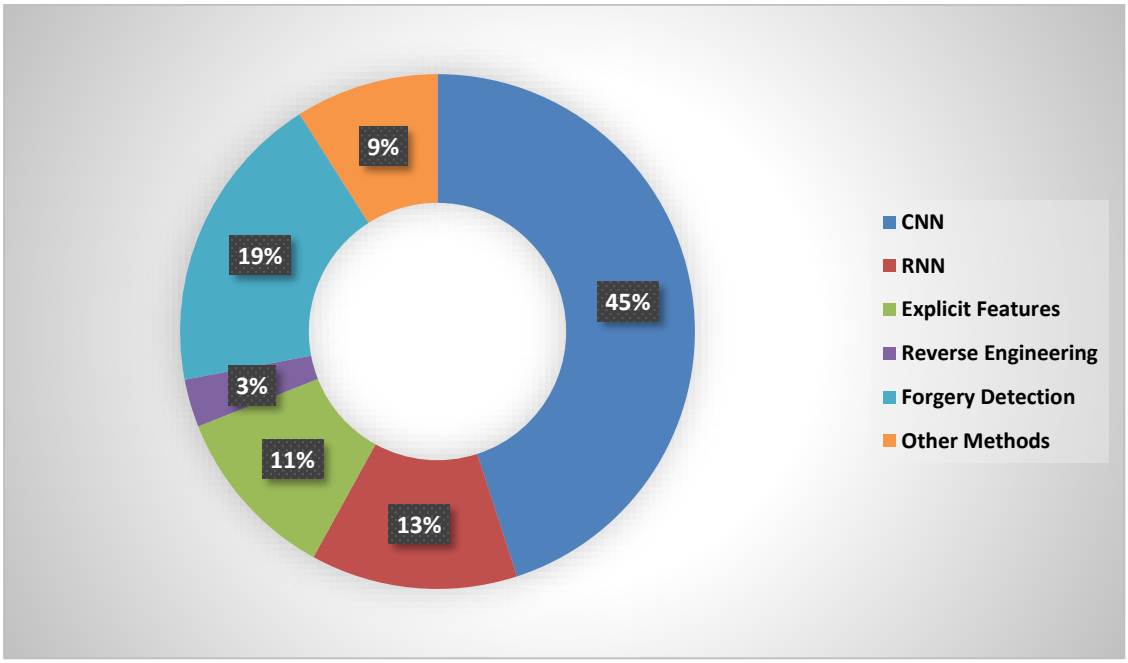


Figure 10: Distribution of reviewed articles representing each method

3.2 RNN

Recurrent Neural Networks analyze sequential inputs like text, image, speech, video, and output in a feedback loop. A network with feedback loops is created, which allows RNNs to retain information and train themselves. RNNs have been utilized in image classification [40], video classification [41], object recognition [42], video annotation [43], time-series prediction [44], anomaly detection [45], sentiment analysis [46], speech recognition and other ML and DL tasks. LSTMs (Long Short-Term Memory networks), a special case of RNNs, have been widely utilized in fake news classification tasks in multiple modalities. These train much faster and can perform complex classification tasks than other RNNs. 13% of the articles reviewed in this paper performing multi-modal fake news detection have utilized RNNs and a combination of RNNs and CNNs.

3.3 Reverse Engineering

A popular method for fake news detection is to look up the content in question on the internet. Post's credibility is assured by matching it with the occurrences that appear in the search results. The methodology used under reverse engineering for fake news detection is Reverse Image Search. Search engines like Google, Yahoo, and Bing allow their users to input a query image and provide them with relevant information about the image. This process is utilized in fact-checking

the authenticity of an image. We can get to know how old an image dates back and where did it appear first. Metadata can also be extracted from such visual data. It also helps us verify the context of the image to the text it accompanies. This method is used by automated fake news detection tools, applications, or web plugins.

Table 1: Comparative analysis of fake news detection methods

Domain	**Method**	**Applications**	**Advantages**	**Disadvantages**
Deep Learning	Convolutional Neural Networks	Image, video, object, speech, time-series, handwritten data recognition, text mining [30-37]	Adaptively learn classification features	Cannot detect image manipulation on its own Require large labelled datasets
Deep Learning	Recurrent Neural Networks	Time, sequence-based predictions, Text, speech, image, video processing, object recognition [40-46]	Great memorizing capacity, usable with CNNs	Slow computation, difficult training, hard to process long sequences, exploding and vanishing gradient
Reverse Engineering	Reverse Image Search	Image origin tracking, metadata tracking	Allows verifying variations in images by matching with originally occurring images and text	Can work with images previously present, cannot detect every type of fake image
Explicit Feature Analysis	Statistical Analysis, Semantic score analysis, User-profile feature analysis, Propagation feature analysis, Geolocation, Psycholinguistics	Explicit feature detection in all modalities based on features present out of the content	Provides prominent non-data features, measures semantics and relevance	Does not include latent features
Data Forensics	Image/video tampering detection, Face manipulation detection	Forged image and video detection, face-swap detection, deepfake detection, editing detection	Easily detects changes made in an image area by editing, removal, addition etc.	Difficult to detect minute manipulations

3.4 Explicit Features

Under explicit features utilized for FND tasks, we categorize statistical features (no. of words, likes, shares, retweets, comments, reactions, etc.), similarity features that analyze the similarities between content and visual information of a news article and state how well both of these are correlated, semantic features that verify meaningfulness of data, user profile features that provide information about users' age groups, backgrounds, faiths and beliefs, inclination, online social behavior, and other relevant profile information, propagation features that help analyze the flow of fake news among networks and people, geolocation features those study areas of fake news generation and propagation and other external features. These features, when combined with other modalities, increase the weightage of detection accuracies. They serve as an important factor for fake news analysis and detection.

3.5 Data Forensics

Images and videos, given the current technological advancements, can be easily edited and tampered with. We have classified fake news detection techniques using forgery detection, splice-detection, copy-move detection, face-swapping detection, face manipulation detection, pixel-based forgery detection, photoshop detection, object-removal detection, repurposing detection, and other similar editing detections under image forensics. This method verifies the credibility of images and videos without a requirement of their original version. Algorithms utilized can detect manipulated regions in images and videos. Popular literature that uses data forensics techniques for image and video classification is explained in section 4.2.

3.6 Other Methods

Few other methods that have been utilized by researchers for fake news detection can be named as co-occurrence matrices, blockchain, pattern recognition, etc. It has become popular to match the semantics between post text, image, and video. Few of the latest works verify if the post's modalities convey the same meaning and then classify them as real or fake. They have provided a new dimension to investigate fake news detection. This area provides opportunities to be explored more, enhance currently available methods, and leverage new ones.

4. Fake News Detection: Trends across Modalities

This section presents an overview of crucial research performed in visual and multi-modal fake news detection. We highlight the vast usage of Deep Neural Networks and forgery detection techniques in multi-modal analysis through the survey. We present the survey classified based on the modalities used for fake news detection.

Qureshi and Deriche [47] explained the taxonomy of types of forgeries found in images: copy-move forgery, image retouching, resampling, image splicing. They have also discussed pixel-based forgery detection methods in images that include contrast enhancement detection, sharpening filtering detection, median filtering detection, resampling detection, post-processing editing detection, copy-move detection, and image splicing detection. Brezeale and Cook [48] provided a survey of existing video classification methods that classify videos using text features, audio features, visual features, and combinations. Boididou et al. [49] have reviewed various methodologies for classifying multimedia data on Twitter that include verifying cues, assessing the source and user credibility, content credibility, image forensics, verifying multimedia use task and have described the verification approaches used, namely UoS-ITI, MCG-ICT, and CERTH-UNITN. Anoop et al. [50] deeply studied fake news detection methods on textual modality, image modality, network modality, temporal modality, and knowledge-based approaches. They have also discussed popular datasets for use. Tolosana et al. [51] reviewed existing face-manipulation detection methodologies. Saini et al. [52] neatly summarized supervised, semi-supervised, and unsupervised multi-modal FND frameworks that include baselines like MVAE and EANN. They have compared state-of-the-art by nicely tabulated data.

4.1 Literature Review

Image tampering has become easier than ever, given the advances in photo editing tools. It is crucial to detect such forgeries to keep a check on fake news data. Fake images accompanied by fake news are categorized as in figure 11. The categories are: tampered/edited images, outdated images used with a later situation, and images out-of-context with the accompanying text. Figures 12, 13, 14, and 15 show examples of fake news images. Approaches utilized to detect these fakes include supervised deep learning techniques that require huge training data. Deep neural networks have been successful in the classification of manipulated images from the originals. Table 2 summarizes all the tasks related to fake news detection that involve visual modalities. It helps to

understand the necessary details of the related works easily. In table 3, a summary of supportive works is provided, which utilize data forensics mechanisms to identify tampered visual data.

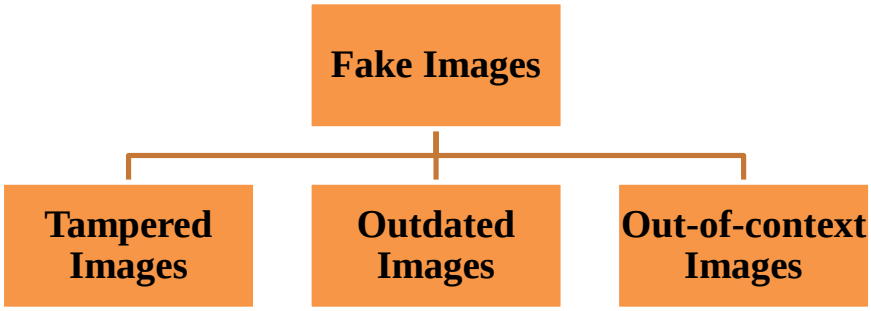


Figure 11: Types of fake images

Elkasrawi et al. [53] allowed users to verify images present online using metadata and feature analysis automatically. The approach has two phases: first involves checking the image for its authenticity, whereas, in second, it is verified if the image has any alterations. The first phase attempts to find other occurrences of an image on the web using Google Image Search. This results in the appearance of the query image in its similar or manipulated versions. Along with several versions of the image, metadata like URL of the image, article in which it appears, publishing date, time, thumbnail, etc., are retrieved. The image's authenticity is validated by matching the timestamps of the query image and search result image. If the image dates back from the date of the query image, it is regarded as fake. In several occurrences, the algorithm uses k-means clustering for the resulting images and their timestamps. In the second phase, image matches are retrieved, and alteration detection is performed. This is done using edge comparison and checking image alignment. The algorithmic features are conglomerated and deployed in a Chrome browser extension for image-based fake news verification.

Wang et al. [54] built a neural network framework for incoming real-time events. This Event Adversarial Neural Network (EANN) framework can handle event-invariant features, thus allowing the detection of fake news on freshly arriving events. With three components in the framework, the first component is a multi-modal feature extractor. Text features are extracted using TextCNN, and a pre-trained VGG-19 is used for visual feature extraction after fine-tuning the hyperparameters. The second component is the fake news classifier, which is built upon the multi-modal feature extraction layers. Classification is performed using a softmax layer for

predictions. The third component is an event discriminator that classifies a post into one of the K events. Its application is to remove event-specific features from the posts and capture only invariant features across all events. EANN is capable of classifying fake news incoming from any type of event.

Figure 12: Original vs. tampered image

Figure 13: Face-swapped images

All other image classification tasks that use deep neural networks utilize CNN models, which have been pre-trained on the ImageNet dataset. Jin et al. [55] highlight that to use CNNs for fake news classification tasks, it is required to train the models on a specific fake news dataset. For image credibility analysis of online news on visual data, they collect an auxiliary dataset of fake images from tweets with 328 K rumor images. These images are labeled in terms of their credibility polarity. A collection of 0.6 M images is used to pre-train a convolutional network with architecture similar to AlexNet. The model is trained using iterative transfer learning. This domain transferred learning algorithm provides favorable results for fake news classification.

Qi et al. [56] are the first to use multi-domain information to classify fake and real images proposing MVNN (Multi-domain Visual Neural Network). They expressed that fake news images are constructed of different physical and semantic features than in real news images. They fused the frequency domain and pixel domain feature extraction sub-networks using an attention mechanism to identify real and fake images. Their proposed network consists of three components: a CNN-based frequency domain sub-network, a pixel domain sub-network built with CNN-RNN to extract semantic features, and a fusion subnetwork. As a whole, their network classifies fake news by using only image features, disregarding linguistics, which is a limitation of the task.

Vishwakarma et al. [57] performed FND by scrapping and authenticating web searches. The work proposes an architecture that enables the verification of text-embedded images. Extraction of text from images is supported by Optical Character Recognition (OCR). The extracted text undergoes Name Entity Recognition steps to obtain named entities mentioned in the

text. These strings are used as queries for Google Search to find matching results. The resulting links are categorized into reliable or unreliable. Next, the entities extracted from images are checked against the search result articles' titles, further classifying them into real or fake.

Pasquini et al. [58] also verified the integrity of online news. They have also demonstrated that fake news detection in visual has a dependency on visual forensics. Their focus is to detect news web pages that contain misleading images. Their framework automatically detects related news articles on the internet with similar images of an event. The detection is performed using metadata features present along with the images and by comparing similar features in the set of related images.

Using image forensics techniques for fake news classification, Huh et al. [59] utilized self-learned consistency to detect image manipulation and splicing. It is done by comparing patches of images. Tampered images show low consistency scores between patches. Models are evaluated for splice detection and splice localization on unannotated images. It is an efficient way to detect fake/manipulated images where copy-move forgery, object addition, or removal have been done.

SAME (Sentiment-Aware Multi-modal Embedding) is a deep end-to-end embedding network that exploits users' sentiments to classify fake news. Cui et al. [60] have incorporated users' sentiments with images, profiles, content, and comments into a multi-modal framework. The model intakes post content, image, and user profile and feeds them to a Multi-Layer Perceptron (MLP), pre-trained VGG-19 network, and another MLP, respectively. The three modalities are fused using an adversarial mechanism.

Jin et al. [61] performed a classification between real and fake tweets using a two-level classification model: topic level and message level. They proposed that tweets among themselves have a strong relationship in terms of event or topic, and tweets clustered in a topic have similar credibility values. It provided a better result than uni-level classification. Tweets with similar images (verifying by features such as resolution, image popularity, etc.) were clustered under the same topic. They extracted topic-level features and message-level content features, user features, and other available features and classified them on both levels. Topic-level classification

Table 2: Summary of crucial work

Reference	Task	Modality	Technique	Classification	Dataset	Accuracy /F1 Score

[61]	Tweet labeling for FND	Text, Image	Two-level classification (Topic and message-level)	Binary	Twitter (Mediaeval 2015[1])	0.94 (F-score)
[62]	Fake news detection	Text, Image	Utilizing user-profile features	Binary	FakeNewsNet	>0.90 (F-score)
[22]	Fake news detection	Text, Image	BERT, VGG19	Binary	Twitter, Weibo	77.77%, 89.23%
[55]	Fake news detection	Image	Image splice detection, splice localization	Binary	Columbia, Carvalho, Realistic Tampering, In The Wild, Hays and Efros	0.91 (mAP)
[65]	Fake tweet and its user identification	Text, Image	Reverse image search, User analysis, crowdsourcing	Multi-class (fake, legitimate, not sure)	Twitter	
[66]	Fake image classification of Hurricane Sandy	Text, Image	Temporal analysis, Naive Bayes, decision tree	Binary	Twitter	97%
[67]	Image tampering detection, text-image coherence detection	Text, Image	Image forensics	Binary	Mediaeval2016, BuzzFeedNews, CrawlerNews	>75%
[68]	Fake news detection	Text, Image	CNN	Binary	TI-CNN	0.92 (F-score)
[110]	Microblogs news verification	Image	Visual and statistical features	Binary	Sina Weibo	83.6%
[56]	Image verification	Image	Metadata, feature analysis	Binary		72.7%, 88%
[111]	Fake news detection	Text, Image	Bi-LSTM, VGG19	Binary	Twitter, Weibo	74.5%, 82.4%
[72]	Video annotation	Video	DCT, HSV, SURF, AOF, ResNet, GoogLeNet		Twitter	0.40 (F-score)
[112]	Video verification	Video	Contextual Cues	Binary	IVC, FVC	0.9 (F-score)

[97]	Fake news detection	Text, Image	VGGNet, sentiment analysis	Binary	PolitiFact, GossipCop	~75, ~80 (Macro, Micro F1)
[92]	Fake news detection	Image	CNN, Iterative Transfer Learning	Binary	Weibo	77%
[113]	Rumour detection on microblogs	Text, Image	LSTM, att-RNN, VGG19	Binary	Twitter, Weibo	78%, 68%
[93]	Fake news detection	Image	CNN (frequency domain), CNN-RNN (pixel domain), Bi-GRU	Binary	Twitter, Weibo	84.6%
[91]	Fake news detection	Text, Image	Text-CNN, VGG19	Binary	Twitter, Weibo	71.5%, 82.7%
[114]	Fake news detection	Text, Image, Source	CNN	Binary	Twitter	82.47% (F-score)
[115]	Semantic Integrity Assessment	Text, Image	MAE, Bi-DNN, VSM	Binary	MAIM, Flickr30, MS COCO	0.75, 0.89, 0.94 (F-scores)
[100]	Fake News Detection	Image	Montage detection (feature-based approach), SIFT, SURF		COCO, INRIA	>90%
[116]	Fact-Checking on image-claim pair	Text, Image	Cosine Similarity, Embedding similarity	Binary	Snopes, Reuters	80.1%
[117]	Fake news detection	Text, Image	Machine Learning Algorithms	Multi-label	Kaggle Fake News Dataset 2017	85.25%
[70]	News consistency verification	Text, Image, Location, Events	CNN		TamperedNews, News400	94%
[69]	Fake News Detection	Text, Image	Text-CNN, Visual, Similarity features	Binary	PolitiFact, GossipCop	87.4%, 83.8%

Results were fused into message-level feature-vector as an extra feature and then trained the classifier. Each tweet was given a separate label instead of labeling each event.

Shu et al. [62] presented a way to utilize user profile features for fake news detection. They extracted and studied explicit and implicit user profile features, also studying which users were most likely to share real and fake news. They have studied users' geolocations, profile images (using CNN, pre-trained VGG16 model), and political bias. PCA was used for dimensionality reduction of profile features. They compared fake news detection performance using UPF (User Profile Features) to multiple approaches, including RST (Rhetorical Structure Theory), LIWC (Linguistic Inquiry and Word Count), RST_UPF, and LIWC_UPF concluding that UPF and UPF-allied techniques provided higher accuracies than others. They also proved that implicit and explicit features, when combined, provided greater results than each being used individually.

Ajao et al. [63] applied hybrid CNN and LSTM-RNN models for text and image classification. LSTM RNN was used to process and classify text sequences. Another model used was LSTM along with dropout regularization 0.2 to remove over-fitting. The third model incorporated a CNN layer after the word-embedding layer of the LSTM model. The models plain vanilla LSTM, LSTM-CNN model, and LSTM with dropout regularization performed in the order of decreasing accuracies 82%, 80%, and 74%, respectively. Under-fitting and lack of sufficient training data account for the low accuracy of the LSTM-drop model. They also showed that hybrid deep learning models' usage provides considerably good accuracy without requiring huge training data.

Singhal et al. [64] have designed a model named SpotFake for detecting multi-modal fake news without any subtasks like finding correlations between textual and visual data. The model consists of a textual feature extractor module that uses BERT (Bidirectional Encoder Representations from Transformers), a visual feature extractor module that uses VGG19, and a fusion module using simple concatenation. The results outperformed state-of-the-art methods EANN and MVAE and provided higher accuracies equal to 77.77% and 89.23% on Twitter and Weibo, respectively.

Figure 14: Out-of-context image (Mumbai's park image accompanied with Hyde Park's news), (Credit: theconversation.com)

Figure 15: Out-of-context image (A lion taken to vet displayed as a lion been strapped to capture MGM logo), (Credit: snopes.com)

Trumper [65] developed a web tool, 'Fake Tweet Buster,' for users to check a tweet's credibility. The user needs to enter a tweet URL, and the application provides a result as fake or legitimate. The tool works on reverse image search (Google Images and TinEye), user analysis, and crowdsourcing. The tool provides the user with matching images, image data (old or new), tweet information, and classification result. It allows the tool-user to enter his opinion about the tweet as fake, legitimate, or unsure. This crowdsourced data can be used in the future to provide value to credit score.

Gupta et al. [66] extracted tweet information and images from Twitter related to Hurricane 'Sandy'. They analyzed that retweets contained many fake images (86%) rather than original tweets. The approach used temporal analysis to study the propagation of fake images over the Twitter network considered a graph with multiple nodes. They created graph networks for followers and retweets of each user studied. For classification, they used user-level features (giving poor results) and tweet level features (providing effective classification), experimenting with two machine learning techniques: Naïve Bayes classifier and J48 Decision Tree classifier, providing 91% and 97% accuracy, respectively.

Lago et al. [67] created a fusion of image forensics algorithms to distinguish between real and fake images. The various approaches used are Image Forensics Methods (Error Level Analysis, Block Artifact Grid Detection, Double Quantization Likelihood Map, Median-filter noise residue inconsistencies detection, JPEG Ghosts, Color Filter Array), Splicebuster, and with the use of CNNs. They also verified the coherence of online news with their accompanying images. Texts were analyzed using TF-IDF or its higher version, STF-IDF. To make classifications, they chose Random Forest and Logistic Regression.

Yang et al. [68] used convolutional networks for both textual and visual fake news detection. They analyzed latent features and explicit features of text and image sub-branches and fused them using early fusion to classify news.

Huckle and White [69] utilized blockchain and cryptography to trace the origination of fake content. Their focus is to determine where the fake information comes from rather than analyzing its structure and features. Their approach lies in determining the cryptographic hash functions of the text and associated image. The verification is made by matching the hashes of the versions of an image occurring at different places. Same images from different occurrences resulting in different hashes indicate that the image has been altered. This principle forms the base of their fake news detection mechanism.

Krishnan and Chen [70] identified tweets containing fake news using data mining, statistical analysis, reverse image search, and cross verification. They have divided their framework into two components: Core and Website. The Core module fetches tweets, extracts the feature set, performs classification, and returns predictions. The Website module is mainly responsible for crowdsourcing and collecting users' guesses about the post's credibility. It also returns the final decision about the post to the end-user. Classification is performed using the J48 decision tree classifier and Support Vector Machine (SVM). The crowdsourced data is stored in a database for future re-training of classification models and performance improvement.

The NewsVallum approach introduced by Armano et al. [71] uses text-image semantics for fake news detection. It focuses on multi-modality using text and image features for classification with deep neural networks and reinforcement learning. It evaluates the credibility of news spread online on a daily or hourly basis.

Zhou et al. [72] exploited the similarity between text and images to detect if a news article is true. The proposed model comprises three modules where the first one is a multimodal feature-extractor, the second module is a classifier, and the third module determines the cross-modal similarity between text and image in the post. They have used a text-based one-dimensional convolutional network for text-classification, and image features are extracted using VGG. Both the feature extractors are combined using a concatenation operation. In the third module, the semantic relation between the text and image pair is calculated using cosine-similarity. This helps in identifying if the modalities of a post express the same meaning. A post is classified as fake or real, depending on all of these features.

Chen et al. [73] argue the need for an automatic fake news detector tool for evaluating the integrity of any news online.

Using multi-modal features, Budack et al. [74] measured the consistency between the modalities for fake news verification. The proposed work evaluates the coherence between text and image data in an unsupervised manner. Textual module extracts persons, entities, and locations using Named-Entity Recognition (NER). POS Tagging is applied, and subsequently, embeddings are calculated using fastText. For visual features, the ResNet model is used for verification. Persons, locations, events, and scene context verification is performed in the cross-modal entity verification process. This model applies to real-world news classification.

Parikh et al. [75] performed the task on tweet text and images providing a web application. The UI allows users to upload a screengrab of a tweet from which the model extracts useful information like tweet text, image, username, timestamp, location, etc., and predict the authenticity of a tweet using these features.

It has become easier to create fake scenarios in videos by replacing, removing, or adding objects, adding text, and swapping face in the recent technological eras. These manipulated videos in which a person's face are swapped by another's are termed as 'deep-fakes.' Such videos have created nasty issues of defamation. Forensics based detection methods majorly detect copy-move manipulations or tampered frames. Neural networks have been applied to detect morphed faces in visual data.

Nixon et al. [76] annotated videos as real or fake to verify the news. They analyzed the text of news stories and videos circulating online and used their metadata to fact-check and annotate the videos. In textual analysis, the author checks the stories for correctness, distinctiveness, homogeneity, and completeness and then groups them under clusters. The videos related to these news stories were then retrieved and annotated based upon fragment information.

Bagade et al. [77] developed a fact-checking web and mobile application 'Kauwa-Kaate' for full-article verification incorporating text, images, and videos. Their proposed system provides a user-friendly interface to query and fact-check information as and when they encounter it. The algorithm scrapes news articles from fact-checked and trusted news sources available on the internet and maintains a repository in the backend. The verification is carried out by matching the

query item with news articles in the repository. Devoting a platform entirely for fact-checking, is a very practical method for users to verify fake news.

Using several tweet-based and user-based features, Boididou et al. [78] have introduced a model that predicts tweets as fake or real depending upon the majority vote from individual classifiers that use different features. The approach has displayed a successful classification of news from a variety of events.

An event rumor detection mechanism for Sina Weibo has been designed by Sun et al. [79]. The model uses many features, which are content-based linguistic features, user-based features, and multi-media features. The model is suitable for detecting rumors in the form of text, image, and video.

4.2 Other Supportive Works

Detecting fake news in visual data is closely related to identifying manipulations in images and videos. The utilization of data manipulation detection algorithms for fake image/video detection offers a worthy scope. In this sub-section, we identify some of the important manipulation detection techniques that could be useful in fake news detection tasks.

A new convolutional layer has been proposed by Bayar and Stamm [80] to classify unaltered and manipulated images. Fake images can be of the misleading type where news content and accompanying image are unrelated or of the tampered type where images are forged to create a fake scenario. Detecting forged images can prove highly beneficial to detect fake news and news containing fake images. The new layer can learn manipulation detection features without needing to extract preliminary features. This layer has been incorporated into a CNN architecture to detect multiple manipulations. For model training, ReLU activation and SGD (Stochastic Gradient Descent) is used. CNNs train faster with ReLU. Binary classification classifies images into unaltered vs. tampered images with 99.31% accuracy. Multi-class classification is used to classify images based on types of forgeries used: median filtering, Gaussian blurring additive white Gaussian noise, re-sampling vs. authentic image with 99.10% accuracy.

Sabir et al. [81] and Jaiswal et al. [82] have detected image repurposing where unaltered images are put together with false metadata in a news item. Pomari et al. [83], Zampoglou et al. [84], and Wu et al. [85] detected fake images by checking if the images or their portions have been sliced.

Table 3: Examples of data manipulation detection techniques

Reference	Task	Modality	Technique	Classification	Dataset	Accuracy /F1 Score
[39]	Deception detection	Text, Video, Audio	3D-CNN, CNN, openSMILE	Binary	Courtroom trials	96.14%
[105]	Tampered video detection	Video, Audio	Speaker inconsistency detection	Binary	VidTIMIT, AMI, GRID	<1% (EER)
[81]	Image repurposing detection	Text, Image, Location	CNN (VGG19), Word2Vec	Binary	MEIR	0.80 (F-score)
[102]	Fake video detection	Video	RNN, CNN (InceptionV3), LSTM	Binary	Deepfake videos, HOHA	>97%
[103]	Forged image and video detection	Image, Video	Capsule Forensics, VGG19	Binary	Deepfake	99.23%
[91]	Face manipulation detection	Video	RNN, CNN (ResNet50, DenseNet)	Binary	Deepfake, Face2Face, FaceSwap	96.9%
[83]	Splice detection	Image	CNN (ResNet50), Illumination Maps	Binary	DSO, DSI, Columbia	96%
[86]	Fake image detection	Image	CNN ensembles	Binary	CelebA, PGGAN	99.99% (AUROC)
[87]	Image translation detection on compressed and uncompressed images	Image	Conventional methods, CNN (DenseNet, InceptionNetV3, XceptionNet)	Binary	CycleGAN	89%
[88]	GANs vs. real image detection	Image	VGG19	Binary	RAISE, Rahmouni	100%
[99]	GANs Image detection	Image	Co-occurrence matrices, CNN	Binary	cycleGAN, starGAN	99.45%, 93.42%
[80]	Image manipulation detection	Image	New convolutional layer	Binary and multi-class	Various camera images	~99.10%

[104]	Face Spoofing Detection	Image	CNN	Binary	CASIA, REPLAY-ATTACK	<5% (HTER)
[118]	Fake image detection	Image	Mixed Adversarial Generators		FantasticReality	0.61 (mAP)
[98]	Fake image detection	Image	Color disparities	Binary	CelebA, LFW, generated images	>90%

Pomari et al. [83] did so by making use of illumination inconsistencies in the image. Computer-generated images and videos have gained huge popularity in the present scenario. It has become fairly easy to generate fabricated content through computers. Such content is entirely unreal or mixed with some kind of real-world entities. The result is a fake product that is not reliable at all. Tariq et al. [86], Marra et al. [87], Nguyen et al. [88], Rahmouni et al. [89], and Rezende et al. [90] have identified fake images generated by computer machines. Nguyen et al. [88] did this using Modular CNN. Sabir et al. [91], Zhang et al. [92], Zhou et al. [93], Dang et al. [94] have detected tampered faces in images, which can be utilized in detecting fake news where faces of celebrities have been swapped to create fake scenarios.

Wu et al. [95] have used supervised learning to trace various types of manipulations like copy-move forgery, object removal, splicing, and other unknown tampering in images. Photoshop is a widely used tool used by content creators to modify visual content. Wang et al. [96] have detected variations created in images using Adobe Photoshop. Wu et al. [97] created BusterNet for copy-move forgery detection. It is a type of forgery where an object from an image is removed from its original location and moved to a different place in the same image. Li et al. [98] used RGB color components in the images to detect changes that occurred due to tampering. Nataraj et al. [99] detected GAN-generated images using co-occurrence matrices. These matrices describe the distribution of co-occurring pixel values or colors. Steinebach et al. [100] recognized image montages for fake news detection. In image montaging, two or more images or their parts are arranged together by cutting, overlapping, pasting, etc., to make a composite image.

Korshunov and Marcel [101] and Guera and Delp [102] addressed facial manipulations in videos where the face of a person is replaced by another. They used deep neural networks for the task. Guera and Delp [102] also contributed with a dataset of 300 deep-fake videos extracted from websites. They classified videos using CNN and LSTM into pristine and deep-fake categories.

CNN has been used for feature extraction from video frames, concatenated and propagated to LSTM for analyzing sequences temporally. This architecture allows detecting fake videos as short as 2 seconds of length. Korshunov and Marcel [101] showed that using static frame features correspond to higher accuracies than using audio-visual analysis.

Nguyen et al. [103] identified forgeries like replay attacks, computer-generated images/videos by building a capsule network with CNN layers. Videos are analyzed at frame level, and the probabilities of fake and real of every frame are averaged to generate results for the video.

Yang et al. [104] detected swapped faces in images and videos using CNN. Korshunov and Marcel [105] performed tampered video detection using inconsistencies between audio and video features. Classifiers used were GMM (Gaussian Mixture Model), SVM, MLP, and LSTM. Krishnamurthy et al. [39] performed deception detection over a small dataset of 121 courtroom videos. They used Text CNN for textual analysis, 3D-CNN for videos, and openSMILE with MLP for audio analysis. They also utilized micro-expression features such as smile, laughter, frown as another modality for deception detection. Data fusion techniques used were Concatenation and Hadamard + Concatenation. Wu et al. [106] and Rosas et al. [107] have also proposed deception detection on videos using real-life trial data.

Li et al. [108] detected tampered faces in videos to detect swapped faces of celebrities by detecting eye-blinking features. Eye blinking is detected using the LRCN (Long-term Recurrent Convolutional Networks) model, measuring how much open an eye is measured concerning frame coordinates. Features have been extracted using the VGG16 convolutional layer and propagated to a sequence learning module that uses LSTM-RNN, which can retain memory. Then, the probability is determined of how much the eye is open or closed. Bestagini et al. [109] detected local tampering in video sequences. This was done by finding duplicated frames in videos and cross-correlating them with Spatio-temporal frame regions.

5. Benchmark Datasets

Lack of suitable multi-modal datasets have, a lot, hampered the progress in the direction of fake news detection. Deep learning algorithms largely depend on huge amounts of training data, which, being meager, has appeared as a big challenge. The maximum number of fake news detection frameworks built to date have been trained upon data extracted from Twitter, Sina Weibo, or some websites. A few of the other small-sized datasets have been generated for image

Table 4: Benchmark Multi-modal datasets

Dataset	Year	Type	Description	Class	Size	Source
MediaEval[1]	2015	Tweet, Image	Tweets related to 11 events (dev set) and 17 events (test set)	Binary	6,225 real tweets 9,596 fake tweets	Topsy, Twitter API
FakeNewsNet[2]	2018	Text, Image	News Content, Social Context, Dynamic Information, article URLs	Binary	6,000 Fake and 18,000 Real	PolitiFact, GossipCop
Fakeddit[3] [119]	2019	Text, Image	Text, Image, Metadata and Comment	Multi-class	8 Lakh samples	Reddit
Newsbag [120]	2020	Text, Image	News articles and images	Binary	15,000 fake and 2 lakhs real	The Onion, The Wall Street Journal
Newsbag++ [120]	2020	Text, Image	Created by Data Augmentation	Binary	3,89,000 Fake and 2,00,000 Real	The Onion, The Wall Street Journal
FVC[4] [112]	2017	Text, Video	Fake Video Corpus for fake video detection	Binary	55 Fake and 49 Real videos	YouTube
Twitter [113]		Text, Image	Text, Image, Social Context	Binary	6,026 Real and 7,898 Fake	Twitter
Sina Weibo [113]		Text, Image	Data extracted from Chinese online social platform	Binary	4,749 Fake and 4,779 Real articles	Sina Weibo
Politifact [60]		Text, Image	News Content and their corresponding images, Retweet Comments	Binary	432 Fake and 624 Real News	Twitter
GossipCop [60]		Text, Image	News Content and their corresponding images, Retweet Comments	Binary	5,323 fake 16,817 real	Magazines, Newspapers and Social Media
CrawlerNews [67]	2017	Text, Image	News articles and images	Binary	2,500 Images	Google News
Mediaeval 2016[5] [112]	2016	Text, Image	Tweets and images from 53 past events	Binary	17,857 Tweets with 10,628 fake and 7,229 Real	Twitter API

[1] https://github.com/MKLab-ITI/image-verification-corpus/tree/master/mediaeval2015
[2] https://github.com/KaiDMML/FakeNewsNet
[3] https://github.com/entitize/fakeddit
[4] https://github.com/MKLab-ITI/fake-video-corpus
[5] https://github.com/MKLab-ITI/image-verification-corpus/tree/master/mediaeval2016

MFN [64]	2018	Text, Image	Event Centric dataset of tweets and corresponding images	Binary	14,000 Tweets and 500 Images 1,154 Real and 1,154 Fake News Articles	Twitter, Snopes and Webhose
TI – CNN[6] [68]	2019	Text, Image	Text, metadata and Image URLs	Binary	20,015 total articles with 8,074 Real and 11,941 Fake	Over 240 Websites
[117]	2019	Text, Image	Metadata, News Articles	Binary	3,568 Fake, 15,915 Real	Multiple Websites
ReCOVery Dataset [121]	2020	Text, Image, Social Information	News Articles, Tweets related to CovID-19.	Binary	2,029 News Articles, 1,40,820 Tweets	Twitter, Multiple Websites
TamperedNews [74]	2020	Text, Image	News articles and images	Binary	72,561 News Articles	BreakingNews Dataset
News400 Dataset [74]		Text, Image	Tweets, Articles	Binary	400	Twitter, Websites
WhatsApp Dataset [8]	2020	Image	Fake images extracted from WhatsApp groups		8,44,000	WhatsApp

analysis, which still is limited in number and not of optimum quality. Video datasets for this task are very rare and contain videos in the count of 100-200. Other video datasets are not completely relevant. There is an urgent demand for good quality multi-modal datasets that would furnish the need of the hour. The advancements in data augmentation or computer-generated data are beginning to contribute towards building datasets. For the time being, we present a piece of tabulated information about available datasets (image, video, and multi-modal) that have been used in the above-reviewed articles for fake news detection and similar tasks (Table 4). We also list out such datasets that contain news article URL or image/video URL. These datasets can be further improved by extracting visual data using web scraping methods.

[6] https://drive.google.com/file/d/0B3e3qZpPtccsMFo5bk9Ib3VCc2c/view

6. Performance Comparison

In this section, we demonstrate the usage of evaluation metrics utilized by fake news detection tasks and compare their performances based on the most utilized metrics, i.e., accuracy and F1-score. The comparison provided here is irrespective of the dataset but highlights each task's features and methods. We determine how the results have been moving all these years and identify prospective detection methods. Performances are displayed for tasks displaying the results achieved by the experiments on datasets they have used. Visual representations are provided for an easy understanding of how a given model performs when they use a specific set of features.

6.1 Evaluation Metrics

This section discusses the various evaluation parameters utilized for fake news classification tasks. We explain the evaluation methods utilized to examine a model's performance using Accuracy, Precision, Recall, F-score, ROC, AUROC, and EER. In figure 16, a confusion matrix is presented that explains the categorization of rightly and wrongly classified items. We refer to the items as news, images, and videos.

CONFUSION MATRIX	*Actually Real*	*Actually Fake*
Predicted Real	TRUE POSITIVE	FALSE POSITIVE
Predicted Fake	FALSE NEGATIVE	TRUE NEGATIVE

Figure 16: Confusion Matrix

True Positive (TP): This includes rightly predicted positive values, i.e., a piece of real news, image or video is classified as real.

True Negative (TN): This includes rightly predicted negative values, i.e., an originally fake item is correctly predicted as fake.

False Positive (FP): This class contains values where a fake item is wrongly predicted as real.

False Negative (FN): This contains real items wrongly predicted as fake by the model.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \qquad True\ Positive\ Rate\ (TPR) = \frac{TP}{TP+FN}$$

$$F1\ Score = \frac{2(Recall \times Precision)}{(Recall+Precision)} \qquad False\ Positive\ Rate\ (FPR) = \frac{FP}{TN+FP}$$

$$Precision = \frac{TP}{TP+FP} \qquad Recall = \frac{TP}{TP+FN}$$

Table 5: Evaluation metrics used in reviewed articles

References	Accuracy	Precision	F1	Recall	AUC	EER	Others
[85], [61], [67], [68], [70], [76], [81]		✓	✓	✓			
[54], [55], [56], [97], [62], [63], [67], [72], [75], [112], [113], [110], [111]	✓	✓	✓	✓			
[80], [53], [102], [116], [96]	✓						
[101]						✓	
[59], [84], [60]							✓
[115]			✓				
[93], [94], [108], [109], [106]					✓		
[89], [74]	✓				✓		
[49]		✓	✓				
[92], [117]	✓	✓	✓	✓	✓		
[95]			✓		✓		
[66]	✓		✓				
[82]	✓		✓		✓		
[100]		✓					

These values allow us to calculate the corresponding accuracy, precision, recall, and F-score. AUC (Area Under Curve) and ROC (Receiver Operating Characteristics), also called AUROC (Area Under Receiver Operating Characteristics), are calculated using TPR and FPR. The ROC curve is plotted with FPR on the x-axis against TPR on the y-axis. AUC is measured as the area under this curve. AUC values range as real values between 0 and 1, with values closest to 1 being

good or correct classification while values closest to 0 being the worst with poor classification. EER, defined as Equal Error Rate, calculates the error that occurred in classification. Table 5 identifies the evaluation metrics used by the reviewed tasks.

6.2 Result Analysis

Several metrics and parameters have been developed to define the functional performance or, in simpler terms, the algorithms' efficiency on giving the desired output of classification of data from a given dataset. Among them and widely utilized and relied upon metrics are Accuracy (in percentage), Precision, Recall, and F1 values, among several others such as AUC and EER. Almost all major research related to Fake News detection utilizes one or more among the former set of four metrics (Accuracy, Precision, Recall, and F-Score).

Hence, we bring forth such metric evaluation summary of the most relevant and pivotal experiments conducted for fake news detection. Figure 17 demonstrates the results of tasks in terms of F1-scores. We observe that the overall performance stays between 80-95% for methods that use textual and visual features combined. The video classification task, which uses the annotation technique, still has a long way to go. In terms of accuracy (Figure 18), we observe that the range of results is between 70-100%, with an average score of 85%. The majority of fake news classification tasks have relied upon deep neural networks. With the changing time, we also notice an inclination towards forensic algorithms for the same. Trends determine that most of the existing approaches have preferred to use deep learning algorithms due to their efficiency, robust nature, feasibility, and accuracy. Most works have preferred to use more than one feature, i.e., using multi-modal data. Thus, depending on more options and the type of fake information posts can offer. The aim is to consider all parameters that form/alters a user’s perception of a piece of information. Convolutional Neural Networks, with maximum usage in the reviewed articles, have displayed eminent classification performance by exploiting the implicit features. They hold the potential to provide better results in future implementations.

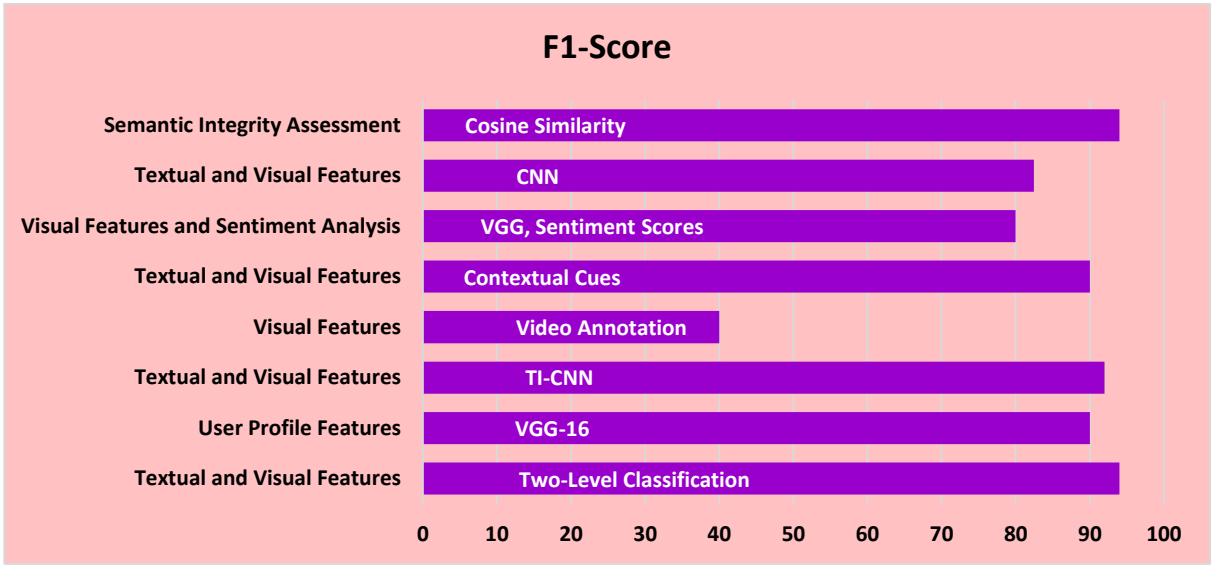


Figure17: Performance comparison based on F1 scores

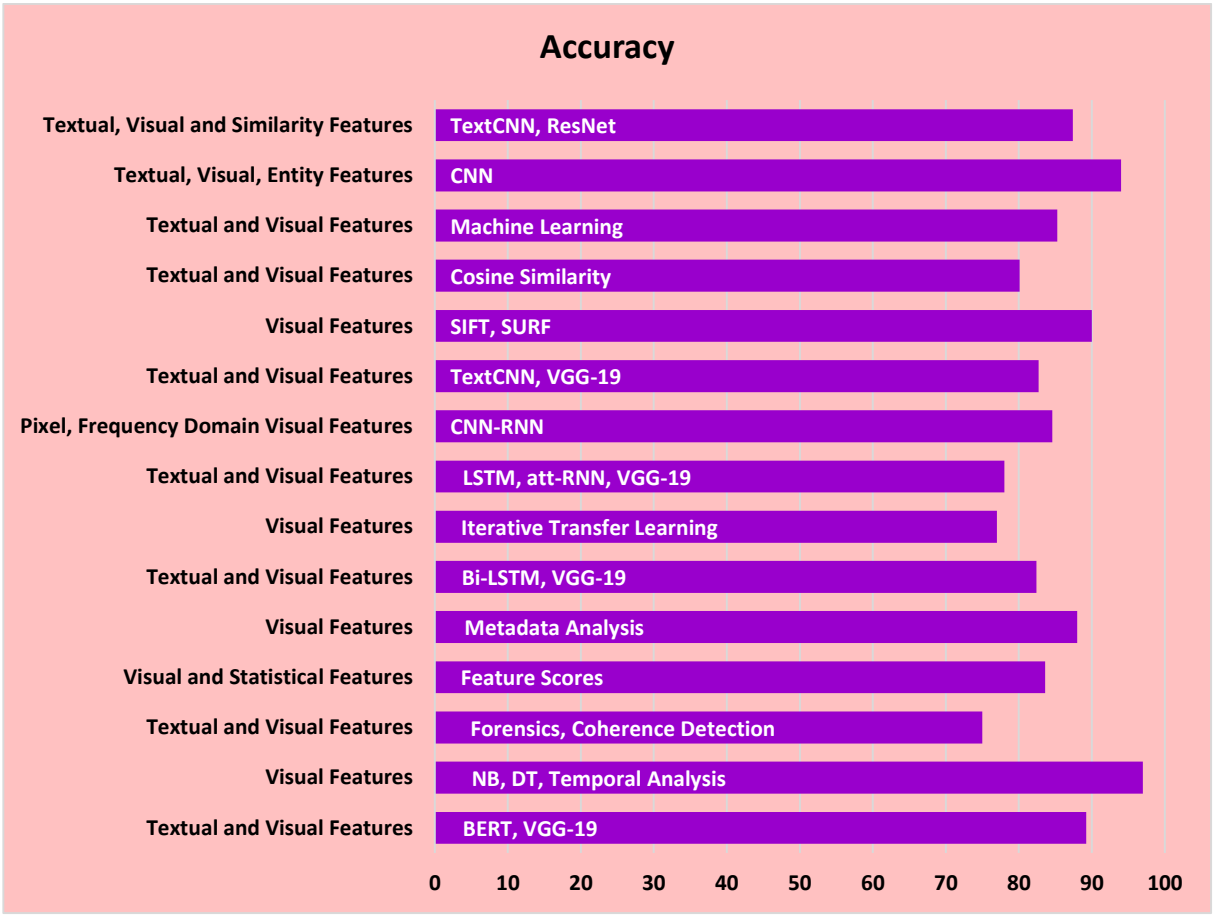


Figure 18: Performance comparison based on accuracy scores

7. Potential Research Directions

Research done in the past years is overwhelming yet, insufficient to cope with the amount of fake news pouring in. Each new happening or event in the world serves as a topic for fake news generation and propagation. In the present scenario, while a pandemic is going on, fake news reaches out to people more swiftly than authentic news is. No data modality is left behind in the race of spreading fake news. Text is not the only type of data of which one should be aware while intaking. People need to be more careful while digesting anything available on the internet because false information could greet us in any form, be it text, image, or video. So is the need for designing efficient and robust detection mechanisms. Analyzing the limitations and research gaps, this section highlights the potential directions where research can be proceeded into.

1. **Datasets:** From the above analysis, it is obvious that there is a shortage of large-scale multimodal datasets. Machine learning and Deep learning algorithms are data-driven. With the shortage of benchmark datasets, it becomes challenging to build detection mechanisms and compare various techniques' performances. Although there are a few text datasets, those with multimodal information are limited and of poor quality. With the advantage of web-scraping mechanisms and free APIs, it has become easier to collect data. To proceed in the direction of multimodality, the collection of large-scale multimodal datasets is promoted.

2. **Real-time Detection:** With the assistance of deep learning algorithms, real-time detection models can be built to use fact-checked articles on the web for training and generate predictions for unseen data. There is a wide opportunity for the development of real-time detectors and automated fact-checkers.

3. **Early Detection:** Fake news detectors are built by feeding past data to algorithms. A baseline comparison is made on previous data. These algorithms are built when the fake information has already spread into the world and affected many. Entrapped within the fake news web, the world requires early detection of false news as and when it appears online. Users can only be benefited from fake news detectors when they provide early detection to prevent the propagation of fake news to a large scale. Early detection would allow intervention and thus mitigation of fake news before it spreads to a larger audience.

4. **Ubiquitous Detection Model:** With many social networking platforms available, it is challenging to incorporate a fake news detection mechanism to separate platforms individually. Similar content makes rounds on multiple platforms because one user can have accounts on various networks. This creates a replica of data on different social networks. With the help of redundant data and manual annotations, classification becomes easy for deep neural networks. A cross-platform system is required that would detect fake content on multiple social platforms. Implementation of models that can train on manually annotated content on one platform and then identify fake news on other platforms is suggested.

5. **Data-oriented Detection:** As far as the previous research is considered, we have very few frameworks that provide credibility assessment to fake content types. Most techniques consider text only, while some allow visual verification. It is challenging for a single system to verify the contents of all data modalities. Such a system would be more beneficial for the general public to authenticate information.

6. **Feature-oriented Detection:** All existing approaches use a limited subset of features, either linguistics, visuals, hybrid, data-centered, sentiment scores, social context, network-based, user-based, or post-based features. These contributing factors of fake news identification could be used all together for dependable predictions.

7. **Integrity Assessment:** In multimodal approaches, existing works perform detection based on features from each type of data independently. In many fake news instances, the post contents are not semantically related. The text, image, or video for a given post could be expressing unrelated context. Few works focus on assessing the semantic integrity of the news. This helps detect false news where data modalities have not been manipulated but are unrelated to each other. Such integrity assessment tools shall help in identifying out-of-context news items.

8. **Embedded Fake News Detection:** Detection of fake embedded content has not been done yet. A large volume of fake news is spreading through such type of data. To cope up with the incoming fake news, this type of detection mechanism is required.

9. **Multilingual Detection:** Current approaches have focused on English language data in text and videos. Due to the spread of fake news through regional languages on the web, multi-lingual approaches should be considered to detect fake news from other languages in the form of text, videos, or embedded content.

10. **Data Manipulation Detection:** With the popularity of image and video forensics techniques, forgery detection in data has become easier. Various manipulation techniques like face-spoofing detection, deepfake identification, tampering detection, splicing, copy-move detection, object removal/addition detection, etc., should be considered and merged with fake news detection mechanisms. There is a need for merging the domains of fake news detection and data manipulation detection.

11. **Browser Plugin/Application Software:** The availability of fake news detector tools in the form of easy-to-use browser plugins, add-ons, software, and mobile applications will enhance their accessibility and serve detection on a user-basis.

8. Conclusion and Future Scope

Uncontrolled and unauthentic data being over-loaded on the web needs appropriate solutions for the complexities being generated and has become a hard nut to crack. Deep learning algorithms are proving efficient and providing effective solutions with remarkable results. These solutions are to be unearthed from unimaginable horizons and that too within a very precise and limited period as the flow of complexities has reached the verge of parallel solutions. There has been a rapid increase in luring solutions for multimodal fake news detection adopting numerous variant techniques. This survey allows us to conclude that deep learning architectures prove astonishingly capable of fake news detection. They have resulted in high accuracies under the text-domain. Recurrent Neural Networks, LSTMs, GRU, Bi-directional GRU have contributed significantly to text classification. When it comes to visual data, Convolutional Neural Networks form the bigger picture. Survey displays that over 40% of methodologies have incorporated CNNs and their combinations with RNNs or other DNNs in their detection frameworks and served brilliant results.

CNNs are taking the lead in computer vision, and allied domains and have become a prospective application for future FND tasks. Many researchers have identified fake images and videos and tampered regions in them, which we review as supportive tasks that can help classify fake news based on fake visuals. We motivate the readers to combine such tasks with FND modules to perform multimodal FND. By fusing modules performing such tasks on different modalities, optimized performances are assured.

There has been the unavailability of symbolic literature in this domain. The progress along the pathways of multimodal fake news detection has been slow. Researchers are unaware of the advancements so far reached. Existing literature is focused upon fake textual news and its detection mechanisms. This survey acknowledges this shortage and provides a broader and intact overview of multimodal fake news detection that incorporates image, video, audio, and their combinations with text. We engage the readers with a taxonomy of detection methods utilized in the referenced articles. We identify the methods with most applications and highlight their prospects. In this review, we have endeavored to cover almost all the significant works performed in relevant domains. We segregate articles based on modalities implied in them. We demonstrate the year-wise trend of work involved and also analyze their method-wise distribution. We neatly summarize all the notable work and highlight important techniques that may pose powerful algorithms in future fake news classification tasks. Further, we described the evaluation metrics adapted in research so far.

We see that accuracy, precision, recall, and F-scores were observed for most of the tasks, whereas some liked to evaluate their models' performances using AUC and ROC. A few other metrics utilized were EER, HTER, TPR, and FPR. Accuracy appears as the most adopted method. Further, owing to the scarcity of multimodal datasets, we regard the obstacles faced in fledged research in the form of not so optimum solutions. Hence, we provide collective information of all the good quality image, video, and multimodal datasets available and previously been used in various tasks. This provides a route for fellow researchers to efficiently choose amongst the few available datasets and perform future research. We also encourage them to build good quality multimodal fake news datasets that would render this domain positively.

Concerning future works, we promote a multimodal framework that would efficiently detect fake news in all forms that revolve around the internet. We suggest exploring the domain incorporating fake news in the form of videos. We also motivate researchers to engage in building versatile multimodal datasets for future use collecting information from websites, online social platforms, and the likes. We encourage the readers to dive deeper into machine learning and deep learning algorithms and fish out ultimate solutions to the problem domain. Our work helps in bridging the research gaps and serve as potential future opportunities to work upon. We conclude this review anticipating that interested researchers will benefit from the information provided and narrow down their interests to this domain to contribute to the society and research community.

References

[1] https://www.news18.com/news/buzz/twitter-trends-with-world-war-iii-after-killing-of-irans-top-commander-in-us-airstrike-2444295.html

[2] https://timesofindia.indiatimes.com/life-style/health-fitness/health-news/coronavirus-myth-vs-fact-whatsapp-forward-claiming-turmeric-and-black-pepper-home-remedy-to-cure-covid-19-is-fake/photostory/76995286.cms?picid=76995474

[3] https://www.pedestrian.tv/news/french-government-cocaine-coke-coronavirus-hoax/

[4] https://en.wikipedia.org/wiki/List_of_unproven_methods_against_COVID-19

[5] https://www.boomlive.in/fake-news/viral-messages-stating-who-protocols-for-lockdown-extension-are-false-7552

[6] https://www.bbc.com/news/world-52224331

[7] https://en.wikipedia.org/wiki/Pizzagate_conspiracy_theory

[8] Reis, J. C., Melo, P., Garimella, K., Almeida, J. M., Eckles, D., & Benevenuto, F. (2020, May). A Dataset of Fact-Checked Images Shared on WhatsApp During the Brazilian and Indian Elections. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 14, pp. 903-908).

[9] A. Gupta, H. Lamba, P. Kumaraguru, and A. Joshi, "Faking sandy: Characterizing and identifying fake images on twitter during hurricane sandy," *WWW 2013 Companion - Proc. 22nd Int. Conf. World Wide Web*, pp. 729–736, 2013.

[10] https://www.business-standard.com/article/current-affairs/gps-chips-and-radioactive-ink-in-new-notes-top-10-fake-news-in-2016-116122600083_1.html

[11] https://en.wikipedia.org/wiki/Fake_news_in_India#Citizenship_(Amendment)_Act_2019

[12] https://en.wikipedia.org/wiki/Fake_news_in_India#Kashmir

[13] Oshikawa, R., Qian, J., & Wang, W. Y. (2018). A survey on natural language processing for fake news detection. *arXiv preprint arXiv:1811.00770*.

[14] Ahmed, H., Traore, I., & Saad, S. (2018). Detecting opinion spams and fake news using text classification. *Security and Privacy*, *1*(1), e9.

[15] Kula, S., Choraś, M., Kozik, R., Ksieniewicz, P., & Woźniak, M. (2020, June). Sentiment analysis for fake news detection by means of neural networks. In *International Conference on Computational Science* (pp. 653-666). Springer, Cham.

[16] Han, Y., Karunasekera, S., & Leckie, C. (2020). Graph Neural Networks with Continual Learning for Fake News Detection from Social Media. *arXiv preprint arXiv:2007.03316*.

[17] https://www.hindustantimes.com/more-lifestyle/covid-19-misinformation-fake-news-on-coronavirus-is-proving-to-be-contagious/story-wmKXCKjESMQIEZoyI9VJ8O.html

[18] https://www.kron4.com/news/media-reports-say-north-korean-dictator-kim-jong-un-dead/

[19] S. Sharma and D. K. Sharma, "Fake News Detection: A long way to go," *2019 4th Int. Conf. Inf. Syst. Comput. Networks, ISCON 2019*, no. May 2020, pp. 816–821, 2019.

[20] Á. Figueira and L. Oliveira, "The current state of fake news: Challenges and opportunities," *Procedia Comput. Sci.*, vol. 121, pp. 817–825, 2017.

[21] F. Torabi Asr and M. Taboada, “Big Data and quality data for fake news and misinformation detection,” *Big Data Soc.*, vol. 6, no. 1, pp. 1–14, 2019.
[22] X. Zhang and A. A. Ghorbani, “An overview of online fake news: Characterization, detection, and discussion,” *Inf. Process. Manag.*, no. February, pp. 1–26, 2019.
[23] E. C. Tandoc, Z. W. Lim, and R. Ling, “Defining ‘Fake News’: A typology of scholarly definitions,” *Digit. Journal.*, vol. 6, no. 2, pp. 137–153, 2018.
[24] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, “Fake News Detection on Social Media,” *ACM SIGKDD Explor. Newsl.*, vol. 19, no. 1, pp. 22–36, 2017.
[25] V. Mosinzova, B. Fabian, T. Ermakova, and A. Baumann, “Fake News, Conspiracies and Myth Debunking in Social Media - A Literature Survey Across Disciplines,” *SSRN Electron. J.*, no. February, 2019.
[26] V. L. Rubin, Y. Chen, and N. J. Conroy, “Deception detection for news: Three types of fakes,” *Proc. Assoc. Inf. Sci. Technol.*, vol. 52, no. 1, pp. 1–4, 2015.
[27] G. Rajendran, B. Chitturi, and P. Poornachandran, “Stance-In-Depth Deep Neural Approach to Stance Classification,” *Procedia Comput. Sci.*, vol. 132, no. Iccids, pp. 1646–1653, 2018.
[28] J. Cao, P. Qi, Q. Sheng, T. Yang, J. Guo, and J. Li, “Exploring the Role of Visual Content in Fake News Detection,” pp. 141–161, 2020.
[29] S. B. Parikh and P. K. Atrey, “Media-Rich Fake News Detection: A Survey,” *Proc. - IEEE 1st Conf. Multimed. Inf. Process. Retrieval, MIPR 2018*, no. April, pp. 436–441, 2018.
[30] Widiastuti, N. I. "Convolution Neural Network for Text Mining and Natural Language Processing." In *IOP Conference Series: Materials Science and Engineering*, vol. 662, no. 5, p. 052010. IOP Publishing, 2019.
[31] Al-Saffar, Ahmed Ali Mohammed, Hai Tao, and Mohammed Ahmed Talab. "Review of deep convolution neural network in image classification." In *2017 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET)*, pp. 26-31. IEEE, 2017.
[32] Karpathy, Andrej, George Toderici, Sanketh Shetty, Thomas Leung, Rahul Sukthankar, and Li Fei-Fei. "Large-scale video classification with convolutional neural networks." In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pp. 1725-1732. 2014.
[33] Singh, Rahul Dev, Ajay Mittal, and Rajesh K. Bhatia. "3D convolutional neural network for object recognition: a review." *Multimedia Tools and Applications* 78, no. 12 (2019): 15951-15995.
[34] Zhao, Bendong, Huanzhang Lu, Shangfeng Chen, Junliang Liu, and Dongya Wu. "Convolutional neural networks for time series classification." *Journal of Systems Engineering and Electronics* 28, no. 1 (2017): 162-169.
[35] Kwon, Donghwoon, Kathiravan Natarajan, Sang C. Suh, Hyunjoo Kim, and Jinoh Kim. "An empirical study on network anomaly detection using convolutional neural networks." In *2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS)*, pp. 1595-1598. IEEE, 2018.
[36] Huang, Jui-Ting, Jinyu Li, and Yifan Gong. "An analysis of convolutional neural networks for speech recognition." In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4989-4993. IEEE, 2015.
[37] Ghosh, Mahmoud M. Abu, and Ashraf Y. Maghari. "A comparative study on handwriting digit recognition using neural networks." In *2017 International Conference on Promising Electronic Technologies (ICPET)*, pp. 77-81. IEEE, 2017.
[38] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” *Artif. Intell. Rev.*, pp. 1–69, 2020.
[39] G. Krishnamurthy, N. Majumder, S. Poria, and E. Cambria, “A Deep Learning Approach for Multi-modal Deception Detection,” 2018.
[40] Mou, Lichao, Pedram Ghamisi, and Xiao Xiang Zhu. "Deep recurrent neural networks for hyperspectral image classification." *IEEE Transactions on Geoscience and Remote Sensing* 55, no. 7 (2017): 3639-3655.
[41] Yue-Hei Ng, Joe, Matthew Hausknecht, Sudheendra Vijayanarasimhan, Oriol Vinyals, Rajat Monga, and George Toderici. "Beyond short snippets: Deep networks for video classification." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4694-4702. 2015.
[42] Liang, Ming, and Xiaolin Hu. "Recurrent convolutional neural network for object recognition." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3367-3375. 2015.
[43] Pan, Pingbo, Zhongwen Xu, Yi Yang, Fei Wu, and Yueting Zhuang. "Hierarchical recurrent neural encoder for video representation with application to captioning." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1029-1038. 2016.
[44] Madhavan, P. G. "Recurrent neural network for time series prediction." In *Proceedings of the 15th Annual International Conference of the IEEE Engineering in Medicine and Biology Societ*, pp. 250-251. IEEE, 1993.

[45] Malhotra, Pankaj, Lovekesh Vig, Gautam Shroff, and Puneet Agarwal. "Long short term memory networks for anomaly detection in time series." In *Proceedings*, vol. 89, pp. 89-94. Presses universitaires de Louvain, 2015.
[46] Ruales, Joaquın. "Recurrent neural networks for sentiment analysis." *IEEE. Colombia: Colombia University* (2011).
[47] M. A. Qureshi and M. Deriche, "A bibliography of pixel-based blind image forgery detection techniques," *Signal Process. Image Commun.*, vol. 39, pp. 46–74, 2015.
[48] D. Brezeale and D. J. Cook, "Automatic video classification: A survey of the literature," *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 38, no. 3, pp. 416–430, 2008.
[49] C. Boididou *et al.*, "Verifying information with multimedia content on twitter: A comparative study of automated approaches," *Multimed. Tools Appl.*, vol. 77, no. 12, pp. 15545–15571, 2018.
[50] K. Anoop, M. P. Gangan, D. P, and V. L. Lajish, *Leveraging Heterogeneous Data for Fake News Detection*. 2019.
[51] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, A. Morales, and J. Ortega-Garcia, "Deepfakes and beyond: A Survey of face manipulation and fake detection," *Inf. Fusion*, vol. 64, pp. 131–148, 2020.
[52] N. Saini, M. Singhal, M. Tanwar, and P. Meel, "Multi-modal, Semi-supervised and Unsupervised web content credibility analysis Frameworks," *Proc. Int. Conf. Intell. Comput. Control Syst. ICICCS 2020*, no. Iciccs, pp. 948–955, 2020.
[53] S. Elkasrawi, A. Dengel, A. Abdelsamad, and S. S. Bukhari, "What You See is What You Get? Automatic Image Verification for Online News Content," *Proc. - 12th IAPR Int. Work. Doc. Anal. Syst. DAS 2016*, pp. 114–119, 2016.
[54] Y. Wang *et al.*, "EANN: Event adversarial neural networks for multi-modal fake news detection," *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, pp. 849–857, 2018.
[55] Z. Jin, J. Cao, J. Luo, and Y. Zhang, "Image Credibility Analysis with Effective Domain Transferred Deep Networks," pp. 1–10, 2016.
[56] P. Qi, J. Cao, T. Yang, J. Guo, and J. Li, "Exploiting multi-domain visual information for fake news detection," *Proc. - IEEE Int. Conf. Data Mining, ICDM*, vol. 2019-Novem, pp. 518–527, 2019.
[57] D. K. Vishwakarma, D. Varshney, and A. Yadav, "Detection and veracity analysis of fake news via scrapping and authenticating the web search," *Cogn. Syst. Res.*, vol. 58, pp. 217–229, 2019.
[58] C. Pasquini, C. Brunetta, A. F. Vinci, V. Conotter, and G. Boato, "TOWARDS THE VERIFICATION OF IMAGE INTEGRITY IN ONLINE NEWS Cecilia Pasquini , Carlo Brunetta , Andrea F . Vinci , Valentina Conotter , Giulia Boato," *Multimed. Expo Work. (ICMEW), 2015 IEEE Int. Conf.*, pp. 1–6, 2015.
[59] L. Self-consistency, M. Huh, A. Liu, A. Owens, A. A. Efros, and U. C. Berkeley, "Fighting Fake News : Image Splice Detection," 2018.
[60] L. Cui, S. Wang, and D. Lee, "Same: Sentiment-aware multi-modal embedding for detecting fake news," *Proc. 2019 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Mining, ASONAM 2019*, pp. 41–48, 2019.
[61] Z. Jin, J. Cao, Y. Zhang, and Y. Zhang, "MCG-ICT at MediaEval 2015: Verifying multimedia use with a two-level classification model," *CEUR Workshop Proc.*, vol. 1436, 2015.
[62] K. Shu, X. Zhou, S. Wang, R. Zafarani, and H. Liu, "The role of user profiles for fake news detection," *Proc. 2019 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Mining, ASONAM 2019*, pp. 436–439, 2019.
[63] O. Ajao, D. Bhowmik, and S. Zargari, "Fake news identification on Twitter with hybrid CNN and RNN models," *ACM Int. Conf. Proceeding Ser.*, no. July, pp. 226–230, 2018.
[64] S. Singhal, R. R. Shah, T. Chakraborty, P. Kumaraguru, and S. Satoh, "SpotFake: A multi-modal framework for fake news detection," *Proc. - 2019 IEEE 5th Int. Conf. Multimed. Big Data, BigMM 2019*, pp. 39–47, 2019.
[65] D. Saez-Trumper, "Fake tweet buster: A webtool to identify users promoting fake news ontwitter," *HT 2014 - Proc. 25th ACM Conf. Hypertext Soc. Media*, pp. 316–317, 2014.
[66] A. Gupta, H. Lamba, P. Kumaraguru, and A. Joshi, "Faking sandy: Characterizing and identifying fake images on twitter during hurricane sandy," *WWW 2013 Companion - Proc. 22nd Int. Conf. World Wide Web*, pp. 729–736, 2013.
[67] F. Lago, Q. T. Phan, G. Boato, and A. Venčkauskas, "Visual and Textual Analysis for Image Trustworthiness Assessment within Online News," *Secur. Commun. Networks*, vol. 2019, 2019.
[68] Y. Yang, L. Zheng, J. Zhang, Q. Cui, Z. Li, and P. S. Yu, "TI-CNN: Convolutional Neural Networks for Fake News Detection," 2018.
[69] S. Huckle and M. White, "Fake News: A Technological Approach to Proving the Origins of Content, Using Blockchains," *Big Data*, vol. 5, no. 4, pp. 356–371, 2017.

[70] S. Knshnan and M. Chen, “Identifying tweets with fake news,” *Proc. - 2018 IEEE 19th Int. Conf. Inf. Reuse Integr. Data Sci. IRI 2018*, vol. 67, pp. 460–464, 2018.
[71] G. Armano *et al.*, “NewsVallum: Semantics-Aware Text and Image Processing for Fake News Detection system,” *CEUR Workshop Proc.*, vol. 2161, 2018.
[72] X. Zhou, J. Wu, and R. Zafarani, “SAFE: Similarity-Aware Multi-modal Fake News Detection,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12085 LNAI, no. 1, pp. 354–367, 2020.
[73] Y. Chen, N. J. Conroy, and V. L. Rubin, “News in an online world: The need for an ‘automatic crap detector,’” *Proc. Assoc. Inf. Sci. Technol.*, vol. 52, no. 1, pp. 1–4, 2015.
[74] E. Müller-Budack, J. Theiner, S. Diering, M. Idahl, and R. Ewerth, “Multi-modal analytics for real-world news using measures of cross-modal entity consistency,” *ICMR 2020 - Proc. 2020 Int. Conf. Multimed. Retr.*, pp. 16–25, 2020.
[75] S. B. Parikh, S. R. Khedia, and P. K. Atrey, “A framework to detect fake tweet images on social media,” *Proc. - 2019 IEEE 5th Int. Conf. Multimed. Big Data, BigMM 2019*, no. September, pp. 104–110, 2019.
[76] L. Nixon, E. Apostolidis, F. Markatopoulou, I. Patras, and V. Mezaris, “Multi-modal Video Annotation for Retrieval and Discovery of Newsworthy Video in a News Verification Scenario,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11295 LNCS, no. Mmm, pp. 143–155, 2019.
[77] A. Bagade *et al.*, “The Kauwa-Kaate fake news detection system: DemO,” *ACM Int. Conf. Proceeding Ser.*, pp. 302–306, 2020.
[78] C. Boididou, S. Papadopoulos, M. Zampoglou, L. Apostolidis, O. Papadopoulou, and Y. Kompatsiaris, “Detection and visualization of misleading content on Twitter,” *Int. J. Multimed. Inf. Retr.*, vol. 7, no. 1, pp. 71–86, 2018.
[79] S. Sun, H. Liu, J. He, and X. Du, “Detecting Event Rumors on Sina Weibo Automatically,” pp. 120–131, 2013.
[80] B. Bayar and M. C. Stamm, “A deep learning approach to universal image manipulation detection using a new convolutional layer,” *IH MMSec 2016 - Proc. 2016 ACM Inf. Hiding Multimed. Secur. Work.*, pp. 5–10, 2016.
[81] E. Sabir, Y. Wu, W. A. Almageed, and P. Natarajan, “Deep multi-modal image-repurposing detection,” *MM 2018 - Proc. 2018 ACM Multimed. Conf.*, vol. 2, pp. 1337–1345, 2018.
[82] A. Jaiswal, Y. Wu, W. Abdalmageed, I. Masi, and P. Natarajan, “Aird: Adversarial learning framework for image repurposing detection,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2019-June, pp. 11322–11331, 2019.
[83] T. Pomari, G. Ruppert, E. Rezende, A. Rocha, and T. Carvalho, “Image Splicing Detection Through Illumination Inconsistencies and Deep Learning,” *Proc. - Int. Conf. Image Process. ICIP*, no. September, pp. 3788–3792, 2018.
[84] M. Zampoglou, S. Papadopoulos, and Y. Kompatsiaris, “DETECTING IMAGE SPLICING IN THE WILD (WEB) Markos Zampoglou , Symeon Papadopoulos , Yiannis Kompatsiaris,” *2015 IEEE Int. Conf. Multimed. Expo Work.*, pp. 1–6.
[85] Y. Wu, W. Abd-Almageed, and P. Natarajan, “Deep matching and validation network: An end-to-end solution to constrained image splicing localization and detection,” *MM 2017 - Proc. 2017 ACM Multimed. Conf.*, pp. 1480–1502, 2017.
[86] S. Tariq, S. Lee, H. Kim, Y. Shin, and S. S. Woo, “Detecting both machine and human created fake face images in the wild,” *Proc. ACM Conf. Comput. Commun. Secur.*, pp. 81–87, 2018.
[87] F. Marra, D. Gragnaniello, D. Cozzolino, and L. Verdoliva, “Detection of GAN-Generated Fake Images over Social Networks,” *Proc. - IEEE 1st Conf. Multimed. Inf. Process. Retrieval, MIPR 2018*, pp. 384–389, 2018.
[88] H. H. Nguyen, N. D. T. Tieu, H. Q. Nguyen-Son, V. Nozick, J. Yamagishi, and I. Echizen, “Modular convolutional neural network for discriminating between computer-generated images and photographic images,” *ACM Int. Conf. Proceeding Ser.*, 2018.
[89] N. Rahmouni *et al.*, “Using Convolution Neural Networks To cite this version : HAL Id : hal-01664590 Distinguishing Computer Graphics from Natural Images Using Convolution Neural Networks,” 2017.
[90] E. R. S. de Rezende, G. C. S. Ruppert, A. Theóphilo, E. K. Tokuda, and T. Carvalho, “Exposing computer generated images by using deep convolutional neural networks,” *Signal Process. Image Commun.*, vol. 66, pp. 113–126, 2018.
[91] E. Sabir, J. Cheng, A. Jaiswal, W. AbdAlmageed, I. Masi, and P. Natarajan, “Recurrent Convolutional Strategies for Face Manipulation Detection in Videos,” pp. 80–87, 2019.
[92] Y. Zhang, L. Zheng, and V. L. L. Thing, “Automated face swapping and its detection,” *2017 IEEE 2nd Int.*

Conf. Signal Image Process. ICSIP 2017, vol. 2017-Janua, no. August, pp. 15–19, 2017.
[93] P. Zhou, X. Han, V. I. Morariu, and L. S. Davis, "Two-Stream Neural Networks for Tampered Face Detection," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 2017-July, pp. 1831–1839, 2017.
[94] L. M. Dang, S. I. Hassan, S. Im, and H. Moon, "Face image manipulation detection based on a convolutional neural network," *Expert Syst. Appl.*, vol. 129, pp. 156–168, 2019.
[95] Y. Wu, W. Abdalmageed, and P. Natarajan, "Mantra-net: Manipulation tracing network for detection and localization of image forgeries with anomalous features," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2019-June, pp. 9535–9544, 2019.
[96] S. Y. Wang, O. Wang, R. Zhang, A. Owens, and A. Efros, "Detecting photoshopped faces by scripting photoshop," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2019-Octob, pp. 10071–10080, 2019.
[97] Y. Wu, W. Abd-Almageed, and P. Natarajan, "BusterNet: Detecting copy-move image forgery with source/target localization," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11210 LNCS, pp. 170–186, 2018.
[98] H. Li, B. Li, S. Tan, and J. Huang, "Identification of deep network generated images using disparities in color components," *Signal Processing*, vol. 174, pp. 1–13, 2020.
[99] L. Nataraj *et al.*, "Detecting GAN generated Fake Images using Co-occurrence Matrices," *IS T Int. Symp. Electron. Imaging Sci. Technol.*, vol. 2019, no. 5, pp. 1–7, 2019.
[100] M. Steinebach, K. Gotkowski, and H. Liu, "Fake news detection by image montage recognition," *J. Cyber Secur. Mobil.*, vol. 9, no. 2, pp. 175–202, 2020.
[101] P. Korshunov and S. Marcel, "DeepFakes: a New Threat to Face Recognition? Assessment and Detection," pp. 1–5, 2018.
[102] D. Guera and E. J. Delp, "Deepfake Video Detection Using Recurrent Neural Networks," *Proc. AVSS 2018 - 2018 15th IEEE Int. Conf. Adv. Video Signal-Based Surveill.*, 2019.
[103] H. H. Nguyen, J. Yamagishi, and I. Echizen, "Capsule-forensics: Using Capsule Networks to Detect Forged Images and Videos," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 2019-May, pp. 2307–2311, 2019.
[104] J. Yang, Z. Lei, and S. Z. Li, "Learn Convolutional Neural Network for Face Anti-Spoofing," 2014.
[105] P. Korshunov and S. Marcel, "Speaker inconsistency detection in tampered video," *Eur. Signal Process. Conf.*, vol. 2018-Septe, no. September, pp. 2375–2379, 2018.
[106] Z. Wu, B. Singh, L. S. Davis, and V. S. Subrahmanian, "Deception detection in videos," *32nd AAAI Conf. Artif. Intell. AAAI 2018*, pp. 1695–1702, 2018.
[107] V. Pérez-Rosas, M. Abouelenien, R. Mihalcea, and M. Burzo, "Deception detection using real-life trial data," *ICMI 2015 - Proc. 2015 ACM Int. Conf. Multi-modal Interact.*, pp. 59–66, 2015.
[108] Y. Li, M. C. Chang, and S. Lyu, "In Ictu Oculi: Exposing AI created fake videos by detecting eye blinking," *10th IEEE Int. Work. Inf. Forensics Secur. WIFS 2018*, 2019.
[109] P. Bestagini, S. Milani, M. Tagliasacchi, and S. Tubaro, "Local tampering detection in video sequences," *2013 IEEE Int. Work. Multimed. Signal Process. MMSP 2013*, pp. 488–493, 2013.
[110] Z. Jin, J. Cao, Y. Zhang, J. Zhou, and Q. Tian, "Novel Visual and Statistical Image Features for Microblogs News Verification," *IEEE Trans. Multimed.*, vol. 19, no. 3, pp. 598–608, 2017.
[111] D. Khattar, M. Gupta, J. S. Goud, and V. Varma, "MvaE: Multi-modal variational autoencoder for fake news detection," *Web Conf. 2019 - Proc. World Wide Web Conf. WWW 2019*, pp. 2915–2921, 2019.
[112] Papadopoulou, O., Zampoglou, M., Papadopoulos, S., & Kompatsiaris, Y. (2017, June). Web video verification using contextual cues. In *Proceedings of the 2nd International Workshop on Multimedia Forensics and Security* (pp. 6-10).
[113] Z. Jin, J. Cao, H. Guo, Y. Zhang, and J. Luo, "Multi-modal fusion with recurrent neural networks for rumor detection on microblogs," *MM 2017 - Proc. 2017 ACM Multimed. Conf.*, pp. 795–816, 2017.
[114] Cédric Maigrot, Vincent Claveau, Ewa Kijak, Ronan Sicre. MediaEval 2016: A multi-modal system for the Verifying Multimedia Use task. MediaEval 2016: "Verfiying Multimedia Use" task, Oct 2016, Hilversum, Netherlands.
[115] A. Jaiswal, E. Sabir, W. Abd Almageed, and P. Natarajan, "Multimedia semantic integrity assessment using joint embedding of images and text," *MM 2017 - Proc. 2017 ACM Multimed. Conf.*, pp. 1465–1471, 2017.
[116] D. Zlatkova, P. Nakov, and I. Koychev, "Fact-checking meets fauxtography: Verifying claims about images," *EMNLP-IJCNLP 2019 - 2019 Conf. Empir. Methods Nat. Lang. Process. 9th Int. Jt. Conf. Nat. Lang. Process. Proc. Conf.*, pp. 2099–2108, 2020.
[117] V. K. Singh, I. Ghosh, and D. Sonagara, "Detecting fake news stories via multi-modal analysis," *J. Assoc. Inf.*

Sci. Technol., no. February, pp. 1–15, 2020.
[118] V. V Kniaz, V. Knyaz, and F. Remondino, "The Point Where Reality Meets Fantasy: Mixed Adversarial Generators for Image Splice Detection," *Adv. Neural Inf. Process. Syst. 32*, no. NeurIPS, pp. 215–226, 2019.
[119] Nakamura, K., Levy, S., & Wang, W. Y. (2019). r/fakeddit: A new multi-modal benchmark dataset for fine-grained fake news detection. *arXiv preprint arXiv:1911.03854*.
[120] Jindal, S., Sood, R., Singh, R., Vatsa, M., & Chakraborty, T. NewsBag: A Multi-modal Benchmark Dataset for Fake News Detection.
[121] Zhou, X., Mulay, A., Ferrara, E., & Zafarani, R. (2020). ReCOVery: A Multi-modal Repository for COVID-19 News Credibility Research. *arXiv preprint arXiv:2006.05557*.

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)
Pune, India. Oct 29-30, 2021

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON) | 978-1-6654-2503-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/SMARTGENCON51891.2021.9645832

Analysis of MISO Converter with Intelligent Control for Solar Water Pumping

Shivangi Yadav
Dept. of Electrical Engineering
Delhi Technological University
Delhi, India
yadavshivangi26@gmail.com

Saurabh Mishra
Dept. of Electrical Engineering
Delhi Technological University
Delhi, India
saurabh.mishra1410@gmail.com

Garima
Dept. of Electrical Engineering
Delhi Technological University
Delhi, India
garima.dtu@gmail.com

***Abstract*—In this study, a multiple-input, single-output (MISO) CUK configuration has been presented for solar powered water pumping system. DC-DC converter contributes to the system's efficiency by assisting in the tracking of maximum power point in conjunction with the MPPT technique. The MISO-CUK converter topology enables a smooth and efficient operation of the system even under fluctuating atmospheric conditions. The MISO-CUK configuration brings about a reduction in the output voltage ripple to a great extent when compared with a conventional single CUK employed to the solar fed water pumping system. The maximum power point is tracked using an intelligent control approach based on an artificial neural network (ANN) . Artificial Neural Networks (ANN) are observed to be more reliant, quick and efficient in cases of dynamic weather conditions. The performance of CUK-MISO DC-DC converter topology implemented along with ANN based MPPT technique for the application of solar fed water pumping is verified in MATLAB/Simulink tool. The suggested system structure portrays a good efficiency and reliable response to the changing irradiation conditions.**

Keywords—CUK-MISO Configuration, SPV Water Pumping System, Maximum Power Point Tracking (MPPT), Artificial Neural Network (ANN), PM-BLDC motor drive.

I. INTRODUCTION

Developing nations are still facing problems when it comes to the electrification of rural, isolated or remote regions. Connecting such areas to the national grid is a much complex problem because of the technological and economic limitations. As a result of this, to cope with the needs of people in these areas, nations are switching to more usage of renewable energy sources like the wind and solar. The renewable sources are abundantly available with limited constraints when compared to the conventional sources of energy. Climate change is another big motivation which has caused the nations across the world, to consider limiting their carbon emissions that are resulting from burning of fossil fuels.

People in rural parts of developing countries such as India rely completely on rainfall to meet their basic water supply needs. There is shortage of drinking water supply, water for irrigating farms etc. In order to solve the problem of water in remote areas solar fed water pumping systems have grabbed attention lately. Solar energy is an abundant, reliable and environment friendly source which is being utilized in innumerable applications as a substitute to non-renewable resources. A generalized solar powered water pumping system constitutes of the PV array module, DC-DC conversion stage, DC-AC stage of inversion, motor and the water pump.

Solar PV arrays produce a low voltage output which needs to be increased to a higher voltage level to meet the requirements of certain applications. Therefore, we require a converter for DC-DC conversion, that is connected to the PV array to increase or decrease its output voltage. The DC-DC converters serves another function, which is to aid in derivation of maximum amount of power from the PV array when it is implemented with a method for tracking of maximum power. A significant number of configurations of DC-DC converters are implemented for on-grid as well as off-grid applications [1-9]. The converters are categorized by their unique topologies like the isolated, non-isolated, interleaved, high-gain etc. In this paper we have suggested a configuration consisting of a CUK converter with multiple inputs and a single output. The advantage of such converters is that, they result in increased extraction of PV power even under mis-matching irradiance levels. CUK converter when interfaced with solar PV systems have shown satisfactory performance exhibiting a good efficiency [10].

The characteristics of solar cells are non-linear in nature. Thus, it has certain limitations as it depends on solar intermittency. It is critical to obtain the most power out of the PV array even at low irradiance levels in order to fully use the PV array's potential. This shortcoming of the solar cells is overcome with the use of Maximum Power Point Tracking (MPPT) methods. MPPT methods is the subject of a lot of study, to ensure the highest efficiency of the solar PV systems. For our system, we have implemented MPPT technique based on artificial neural networks (ANN). Intelligent control techniques provide a control of duty ratio which is independent of the step size. They have a faster response in dynamic conditions when compared to traditional MPPT methods [11]. Numerous other methods for tracking of maximum power point have been discussed in [12-15].

In the PV system, motors are utilized to drive the water pump, which are linked to their shafts.

Permanent Magnet Brushless DC Motor is being preferred now for solar powered water pumping systems owing to their more reliable and efficient operation due to the lack of brush and commutator contacts. They also have simpler control techniques which make them even more favorable over other type of motors. BLDC motors receive

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:00:56 UTC from IEEE Xplore. Restrictions apply.

their input supply from the three phase voltage source inverters (VSI) which in turn is fed by the DC-DC converter. Electronic Commutation in BLDC motor is responsible for providing switching pulses for the inverter. MATLAB/Simulink software is used to observe the response of the entire solar-powered water pumping system.

The following is a description of the paper's format: Section II explains the system's structure, whereas Section III describes the component design process. Section IV presents the control technique. Section V discusses the simulation findings, followed by Section VI, which concludes.

II. SYSTEM STRUCTURE

Solar fed water pumping systems comprise of majorly four components. PV array is the source of power for the complete system as it is responsible for change of solar energy to electrical energy. However, the PV array is only capable of generating a small value of output voltage which needs to be increased to higher level voltages for most applications. For this purpose, DC-DC conversion stage is introduced. These converters not only help in boosting the voltage to the required value but they are also responsible for getting the maximum power from the array, they are connected to for boosting the efficiency of the system as well. DC-DC converters are implemented along with maximum power point tracking techniques so that highest efficiency can be attained. These converters are followed by the DC-AC inversion stage. Three phase inverters are used in the system which obtain their switching pulses with the help of the electronic commutation process. BLDC motors are employed in the system whose shafts are connected to the load which is the centrifugal pump for the purpose of pumping water. For our system the duty ratio for the DC-DC converters is obtained by using Artificial Neural Network (ANN) MPPT technique. Fig.1 demonstrates the overall system configuration.

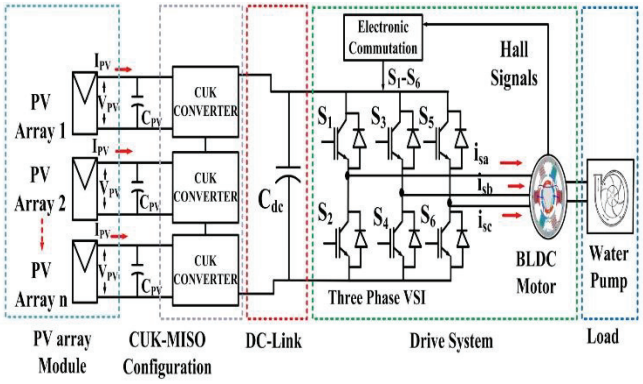


Fig.1 Structure of the System [6]

III. DESIGNING OF SYSTEM COMPONENTS

The next sub-section goes through the design of the various components of solar-powered water pumping system in further detail.

A. PV Array

The system in our study consists of four solar PV arrays which have been connected to the CUK converter. The power capacity of each module is 750 W and thereby the entire PV system is of 3kW. The PV module has a rating slightly above that of the BLDC motor for making up for the losses
which occur in the DC-DC conversion stage. Three PV modules having specifications as given below are series connected and one PV module is connected in parallel to make a single PV array. A total of four such PV arrays are connected at the input side. Table I mentions the PV array's specifications.

TABLE.I SPECIFICATIONS OF PV ARRAY

DESCRIPTION	VALUES
Power at MPP	250 W
Open-Circuit (OC) Voltage	36.4 V
Short-Circuit (SC)Current	8.7 A
Voltage at P_{max}	30.7 V
Current at P_{max}	8.2 A
Cells in Parallel	1
Cells in Series	3

Fig. 2 shows the waveforms of characteristics of the array for standard weather conditions.

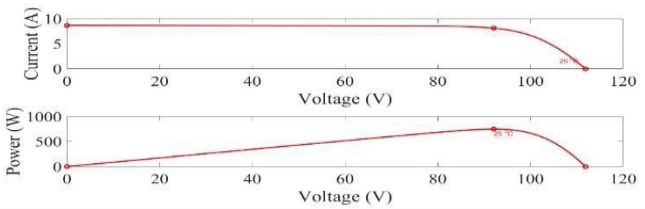


Fig.2 Characteristic Waveforms of PV array

B. DC-DC Converter

The multiple input single output CUK converter is used for DC-DC conversion in the system. Single CUK converter when added as an intermediate stage in the system was seen to produce more ripples in the output voltage as compared to the MISO-CUK converter results. Reduction in the output voltage pulsations is advantageous for the smooth and efficient operation of the load. Fig. 3 shows the circuit diagram of single CUK converter-based water pumping system. Fig.4 shows the structure of CUK converter.

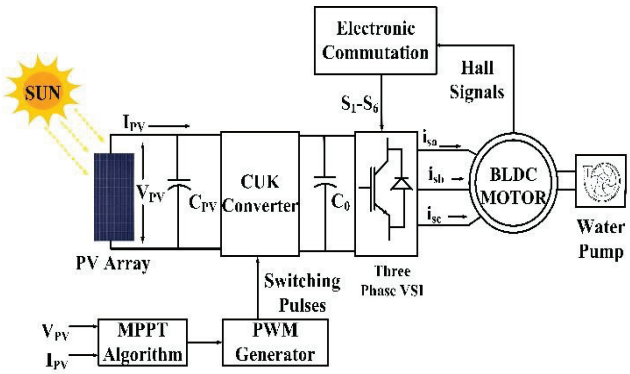


Fig.3 Conventional Single CUK Configuration for Solar Fed Water Pumping System [10]

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:00:56 UTC from IEEE Xplore. Restrictions apply.

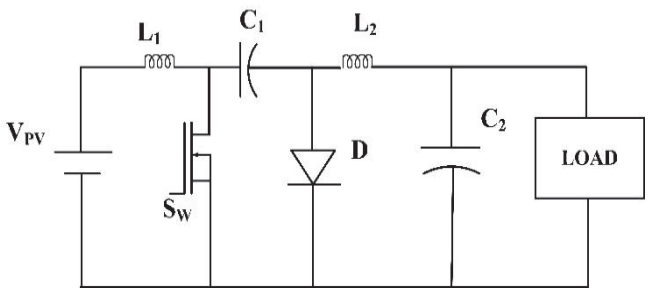


Fig.4 CUK Converter Structure [10]

The equations for determining the components of the CUK converter are given below.

$$P_0 = 750 \text{ W} \quad (1)$$

$$V_{in} = 90 \text{ V} \quad (2)$$

$$V_0 = 150 \text{ V} \quad (3)$$

$$d = \frac{V_0}{V_0 + V_{in}} = \frac{150}{90+150} = 0.62 \quad (4)$$

$$I_{L_1} = \frac{P_0}{V_{in}} = \frac{750}{90} = 8.33A \quad (5)$$

$$\Delta I_{L_1} = 40\% \text{ of } I_{L_1} = 3.332A \quad (6)$$

$$I_{L_2} = \frac{P_0}{V_0} = \frac{750}{150} = 5A \quad (7)$$

$$\Delta I_{L_2} = 40\% \text{ of } I_{L_2} = 2A \quad (8)$$

$$V_{C_1} = V_{in} - V_0 = 90+150 = 240V \quad (9)$$

$$\Delta V_{C_1} = 5\% \text{ of } V_{C_1} = 12V \quad (10)$$

$$\frac{\Delta V_0}{V_0} = 2\% \quad (11)$$

$$L_1 = \frac{V_{in} * d}{f * \Delta I_{L_1}} = \frac{90*0.62}{5000*3.332} = 3.4\ mH \quad (12)$$

$$L_2 = \frac{V_{in} * d}{f * \Delta I_{L_2}} = \frac{90*0.62}{5000*2} = 5.67\ mH \quad (13)$$

$$C_1 = \frac{V_0 * d}{R * f * \Delta V_{C_1}} = \frac{150*0.62}{30*5000*12} = 52.5\ \mu F \quad (14)$$

$$C_2 = \frac{(1-d)}{\frac{\Delta V_0}{V_0} * 8 * L_2 * f^2} = \frac{(1-0.62)}{0.02*8*5.67*25*10^3} = 16.32\ \mu H \quad (15)$$

Where P_0 is the power at output, V_{in} is supplied voltage, V_0 is voltage at output, d is the duty ratio, I_{L1} is the inductor(L_1) current, I_{L2} is inductor(L_2) current, ΔI_{L1} and ΔI_{L2} are the permitted current ripples, f is the frequency, L_1, L_2, C_1, C_2 are the values of the components of the four CUK converters employed in CUK-MISO configuration.

C. Permanent Magnet Brushless DC Motor(PM-BLDC)

There is a wide variety of motors that are being utilized now days for the purpose of solar fed water pumping systems. Brushless DC motor has begun to grab attention because of its advantages over other motors. PM-BLDC motor lacks the brushes and the commutators, which reduces the need of regular maintenance to a great extent. These motors are efficient and reliable and also have simpler control structures, which makes them a preferred substitute of other motors. The electronic commutation process of these motors is responsible for providing the switching pulses of the voltage source inverters (VSI).

D. Centrifugal Pump

With the aid of the BLDC shaft, the centrifugal pump is connected to the water pumping system driven by solar energy. The purpose of the pump is to extract the groundwater. The centrifugal rating chosen for our system is 2.5 kW. The proportionality constant of the water pump is inversely proportional to the cube of speed. To find out the proportionality constant equation given below is used.

$$K = \frac{P_0}{\omega^3} = \frac{2500}{(2*\pi*\frac{3000}{60})^3} = 8.05*10^{-5} \quad (16)$$

IV. CONTROL TECHNIQUE

Two of the components of the solar-powered water pumping system must be controlled. One is the control at the DC-DC conversion stage, by providing it with the appropriate duty cycle for extraction of the highest power possible and another component to be controlled is the voltage source inverter. In the sub-sections below, the control technique for the two separate stages of the system has been detailed.

A. MPPT Using ANN

Innumerable techniques are present today for tracking of maximum power in a solar PV system. Some of the well-known and commonly implemented algorithms and strategies are the Fractional open-circuit voltage, Fractional short-circuit current, Perturb and Observe(P&O) etc. Traditional MPPT approaches, on the other hand, have been found to be dependable and fail to perform efficiently when the weather conditions change constantly. Thus, researchers have now switched to intelligent control methods for finding the maximum power point at different instances. Artificial Neural Networks are capable of training themselves with the variation in atmospheric conditions and are therefore are more reliable and efficient. Fig. 5 demonstrates the basic structure of ANN.

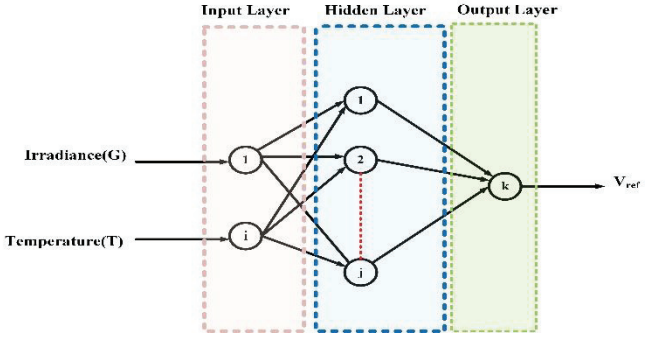


Fig.5 Artificial Neural Network Structure [11]

For our study, ANN based MPPT has been used which takes Irradiance (G) and Temperature(T) as the input. The data generated with the help of these inputs is used to train the artificial neural network using Levenberg-Marquardt algorithm. The trained neural network produces the reference voltage as the output and later it is compared with the PV array's instantaneous voltage. PI controller is fed with the difference between the two voltages. PI controller's output provides the required duty ratio for the DC-DC converter. The artificial neural network is trained using 2500 data points with the help of Levenberg-Marquardt algorithm. Fig.6 demonstrates the process of implementation of the ANN based MPPT method.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:00:56 UTC from IEEE Xplore. Restrictions apply.

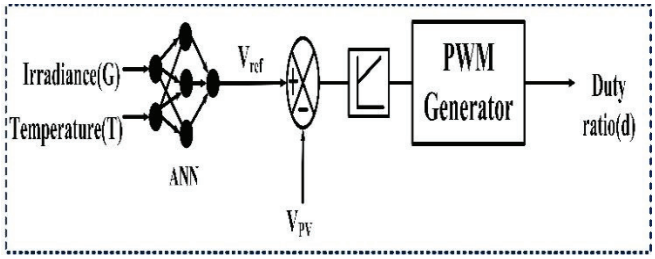


Fig.6 Implementation of ANN based MPPT [14]

B. Electronic Commutaion of BLDC Motor

The mechanism of commutation taking place electronically in the PM-BLDC motor determines the switching pulses of the voltage source inverter (VSI). In this process, the winding currents are commutated with the help of a decoder logic. There are inbuilt sensors which sense the position of the rotor and thereafter the encoder present, utilizes them to generate the hall-effect signals. At a 60^0 interval, a precise set of hall-effect signals is created for every specified series of rotor positions. These encoded signals are then decoded to provide ON/OFF pulses to the inverter. Two switches of the VSI are turned ON at one time whereas the rest of the switches remain in OFF position. This causes the operation of VSI in 120^0 conduction mode. Unlike other motors where moving parts are involved to perform commutation process, electronic commutation does not involve the movement of brushes or commutators. Hence, this process reduces the losses occurring due to conduction. Also, the switching losses caused due to high frequency PWM switching are also eliminated.

V. SIMULATION RESULTS

CUK-MISO DC-DC Converter configuration is implemented with solar powered BLDC motor drive using the MATLAB/Simulink software. This system is subjected to a standard temperature of 25^0 but a varying irradiance within a time span of 2 s. The response of each and every component of the system has been recorded and given in the Fig.7 to Fig.14 below. Some of the inferences that can be drawn from the simulation results are:

- Fig.7 presents the waveforms related to the PV array in reference to the varying irradiance values. The voltage as well as current waveforms of the PV array show variation at time instances t=0.8 s and t=1.2 s, where there is simultaneous transition in the amount of irradiation on the PV array.
- Fig.8 and Fig.9 presents the waveforms related to the inductive and capacitive components of the CUK converter which also show transition with changing irradiance. This proves that the converter and MPPT are working efficiently to find the maximum power point at changing levels of irradiance.
- Fig.10 shows the variation of the voltage and current at DC side of the system. Rated voltage and current are achieved at maximum level of insolation which is $1000W/m^2$.
- Fig.11 depicts the power relationship between the CUK converter's input and output sides. The DC power obtained is the power obtained from all four CUK converters working in conjunction. However, the power on the input side is shown for a single PV array. Rated power is achieved on both input and output sides for highest irradiance levels. With reduction in irradiance level, the value of P_{max} also changes which can be observed in the waveform.
- The responses of the drive system are shown in Fig. 12 and Fig.13. The motor runs at its rated speed for highest insolation.
- Fig.14 shows the transition of duty cycle as MPPT strives to achieve maximum power point. With every change in irradiance level the duty cycle adjusts to a new value to extract maximum power from the array.

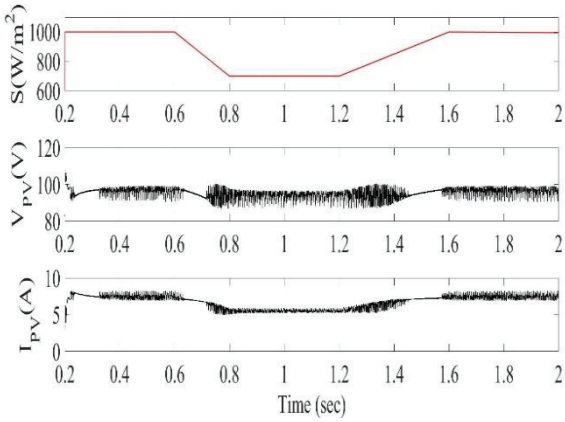


Fig.7 Response of the PV array Voltage and Current with varying irradiance

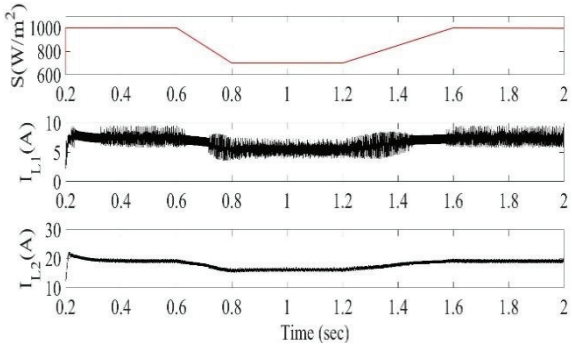


Fig.8 Response of Inductor L_1, L_2 of CUK Converter with varying irradiance

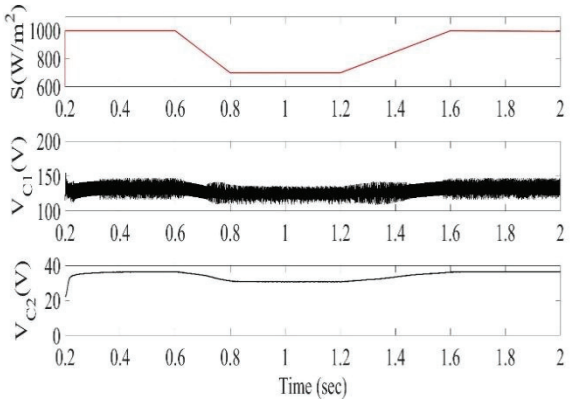


Fig.9 Response of Capacitor C_1 and C_2 of CUK Converter with varying irradiance

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:00:56 UTC from IEEE Xplore. Restrictions apply.

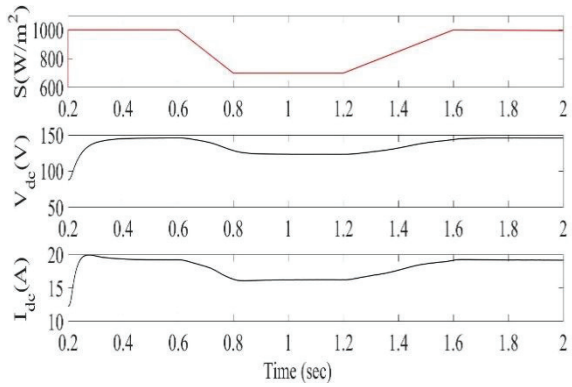


Fig.10 Voltage and Current Output of the DC side with varying irradiance

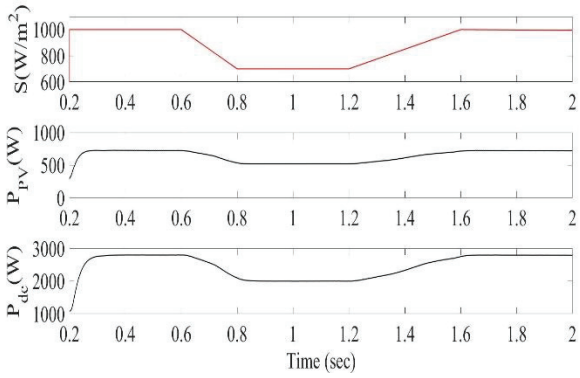


Fig.11 Power curves of one PV array and DC side with varying irradiance

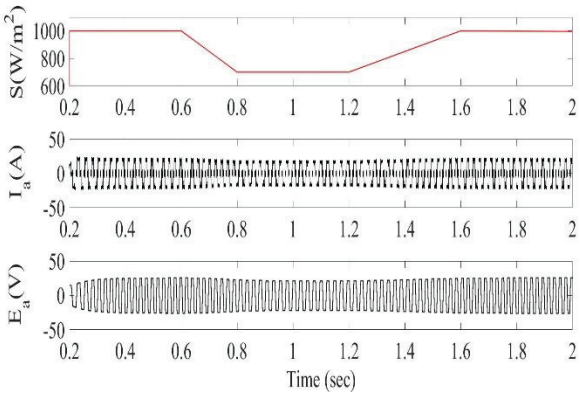


Fig.12 Stator Current and Back-Emf waveforms with varying irradiance

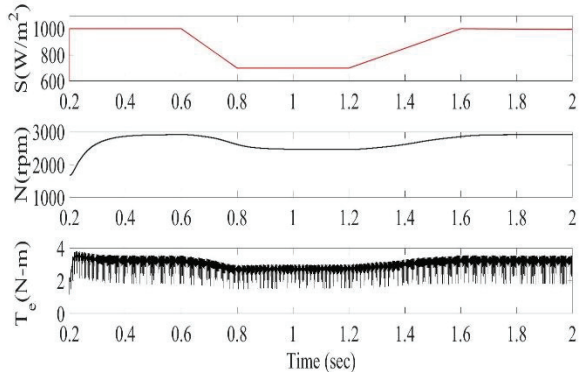


Fig.13 Speed and Torque Variation with Varying Irradiance

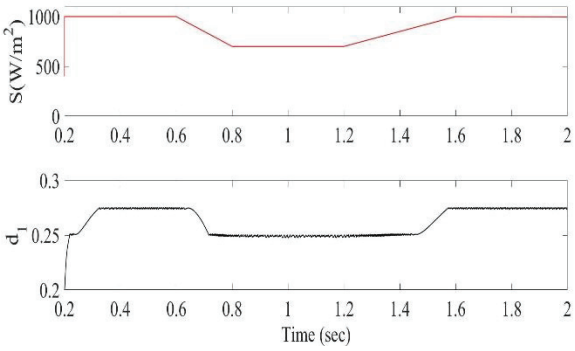


Fig.14 Variation of duty cycle with varying irradiance levels

Comparison in the performance of a single CUK DC-DC converter used with the solar powered water pumping system has been carried out with the CUK-MISO configuration implemented to the solar fed water pumping system in the Fig.15 to Fig18.

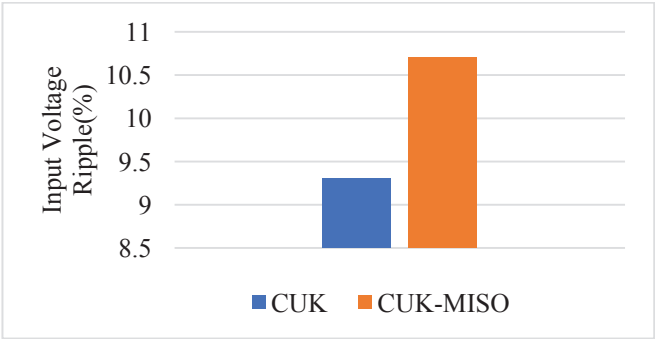


Fig.15 Comparison of Output Voltage Ripple

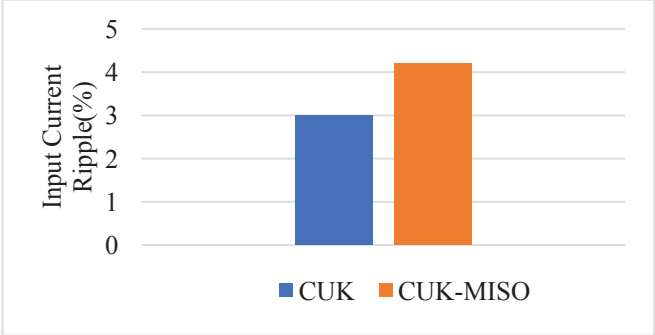


Fig.16 Comparison of Input Current Ripple

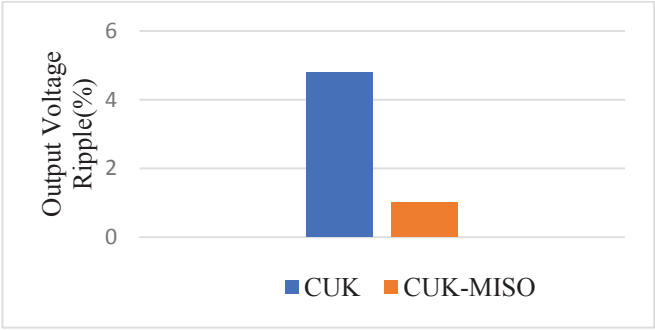


Fig.17 Comparison of Output Voltage Ripple

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:00:56 UTC from IEEE Xplore. Restrictions apply.

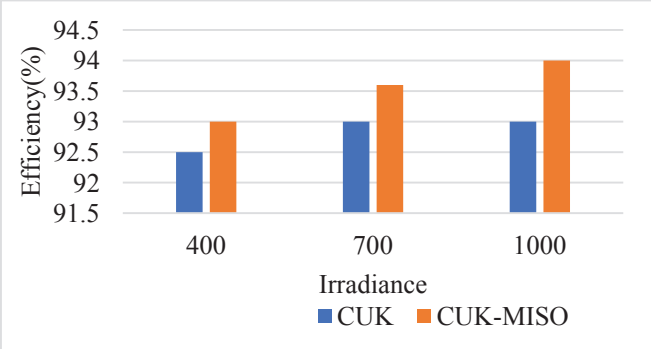


Fig. 18 Efficiency Comparison of CUK and CUK-MISO Configuration at three different irradiance levels

Comparative analysis of the conventional CUK converter interfaced to the PV array and the suggested CUK-MISO configuration presented above shows that with the implementation of the suggested DC-DC converter configuration, ripples in output voltage has been greatly reduced which ensures a smooth operation at the load end. CUK-MISO DC-DC converter also presents a high efficiency value at different irradiance levels. However, number of components in the system is not optimized in this study in order to achieve system simplicity. For this purpose, of simplifying the system other topologies of DC-DC converters can also be studied and researched.

VI. CONCLUSION

For decreasing the output voltage ripples and also to get most power out of the PV array, a multiple input and single output (MISO) CUK converter is suggested in this study. The MISO configuration of CUK converter helps in minimizing the output voltage ripple giving a pulsation free output voltage. The MISO configuration works efficiently in conjunction with the artificial neural network (ANN) based MPPT technique even under dynamic irradiance values. It is observed by the waveforms given in the results that ANN can track maximum power point with greater efficiency in comparison to the conventional MPPT tracking methods. Comparative study is also done between the conventional single CUK converter interfaced to the solar fed water pumping system and the proposed CUK-MISO configuration and the system has shown satisfactory response. The operational efficacy of the PM-BLDC motor solar water pumping system with CUK-MISO configuration and ANN based MPPT technique is supported by simulation findings in MATLAB/Simulink software.

ACKNOWLEDGMENT

Authors want to express their sincere thankfulness to Delhi Technological University (DTU) for supporting our works and also to SMARTGEN-2021 for providing a platform to share our research findings.

REFERENCES

[1] H. Yang, K. Chao, X. Sun, S. An, B. Ren and Q. Zhang, "Analysis, modeling and control of a non-isolated four-port dc-dc converter for independent photovoltaic power generation system," *IECON 43rd Annual Conference of the IEEE Industrial Electronics Society*, pp. 7731-7735, 2017.

[2] P. S. Prakash, R. Kalpana and B. Singh, "Solar Photovoltaic Array fed Push-Pull Buck DC-DC Converter for Telecom Load," *14th IEEE India Council International Conference (INDICON),* pp. 1-4, 2017.

[3] A. Ganjavi, H. Ghoreishy and A. A. Ahmad, "A Novel Single-Input Dual-Output Three-Level DC–DC Converter," *in IEEE Transactions on Industrial Electronics,* vol. 65, no. 10, pp. 8101-8111, Oct. 2018.

[4] O. Abdel-Rahim and H. Wang, "A new high gain DC-DC converter with model-predictive-control based MPPT technique for photovoltaic systems," *in CPSS Transactions on Power Electronics and Applications,* vol. 5, no. 2, pp. 191-200, June 2020.

[5] A. Tomar, P. H. Nguyen and S. Mishra, "SEPIC-MISO Converter Based PV Water Pumping System- An Improved Performance Under Mismatching Conditions," *IEEE 9th Power India International Conference (PIICON),* pp. 1-5, 2020.

[6] Smrity, A. Kumar, N. Bajoria and A. Alam, "Multiple-Input-Single-Output Converter for Hybrid Renewable Sources," *International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN)*, pp. 1-5, 2019.

[7] A. Tomar and S. Mishra, "Multi-input single-output DC-DC converter based PV water pumping system*," IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES)*, pp. 1-5, 2016.

[8] R. Kumar and B. Singh, "BLDC Motor-Driven Solar PV Array-Fed Water Pumping System Employing Zeta Converter," *in IEEE Transactions on Industry Applications*, vol. 52, no. 3, pp. 2315-2322, May-June 2016.

[9] R. Kumar and B. Singh, "Solar PV array fed cuk converter-VSI controlled BLDC motor drive for water pumping," *6th IEEE Power India International Conference (PIICON)*, pp. 1-7, 2014.

[10] A. Kchaou, A. Naamane, Y. Koubaa and N. K. M'Sirdi, "Comparative study of different MPPT techniques for a stand-alone PV system," *17th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, pp. 629-634,2016.

[11] A. Ali et al., "Investigation of MPPT Techniques Under Uniform and Non-Uniform Solar Irradiation Condition–A Retrospection*," in IEEE Access,* vol. 8, pp. 127368-127392, 2020.

[12] R. B. Bollipo, S. Mikkili and P. K. Bonthagorla, "Hybrid, optimal, intelligent and classical PV MPPT techniques: A review*," in CSEE Journal of Power and Energy Systems*, vol. 7, no. 1, pp. 9-33, Jan. 2021.

[13] P. Jena, R. Pudur, P. K. Ray and A. Mohanty, "ANN Based MPPT Applied To Solar Powered Water Pumping System Using BLDC Motor," *IEEE International Conference on Sustainable Energy Technologies and Systems (ICSETS)*, pp. 200-205, 2019.

[14] K. Y. Yap, C. R. Sarimuthu and J. M. -Y. Lim, "Artificial Intelligence Based MPPT Techniques for Solar Power System: A review*," in Journal of Modern Power Systems and Clean Energy,* vol. 8, no. 6, pp. 1043-1059, November 2020.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:00:56 UTC from IEEE Xplore. Restrictions apply.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:10/October-2021 Impact Factor- 6.752 www.irjmets.com

ANALYSIS ON FLOATING SOLAR POWER PLANT

Dr. Jp Kesari*1, Abhay Pratap Singh Shishodia*2, Eshaan Gupta*3

*1Professor, Department Of Mechanical Engineering, Delhi Technological University, Bawana Road, New Delhi-110042, India.

*2Undergraduate Student, Department Of Applied Physics, Delhi Technological University, Bawana Road, New Delhi-110042, India

*3Undergraduate Student, Department Of Applied Physics, Delhi Technological University, Bawana Road, New Delhi-110042, India.

ABSTRACT

This paper focuses on survey of the current drifting sunlight-based plants worldwide as for their ability. Drifting sunlight-based plants can save the region for age. Restrictions to such power plant are land accessibility, land advancement and land securing, substation limits, departure additionally ideal clearances for the venture ashore and departure. These are obstacles for finishing of the undertaking. The greater part of the areas projected by the public authority considering sun powered radiation information in the nation are blistering and dry districts. However, at these areas the radiation seemed, by all accounts, to be higher, the energy yield of these focuses is less because of warming of the sun powered chargers and higher temperature of the outer layer of sun-oriented cells. To beat these issues a creative thought has come in front for establishment of sun-based power plants on the water that is trench tops, water bodies, lakes, dam backwater and repositories, which for the most part has a place with the public authority. This paper uncovers survey with respect to the drifting sunlight-based PV power plants introduced on the planet.

I. INTRODUCTION

One of the biggest problems in India is the energy crisis. Major generation of Electricity is accomplished by coal in the thermal power plants which is approximately 75% of the total power generation. India's total power generation in Dec 2020 was 103.66 billion units, as stated by the data realized by the Central Electricity Authority. These fossil fuels are depleting today because of such extensive use. The limited fossil fuel resources and increasing energy demands are prompting us to move forward towards harnessing solar energy, which is a cost-free, efficient, and never-ending source of energy as well as ecological and sustainable to the environment.

India is the seventh-largest country in the world in terms of area and thus it gets a really good amount of sunshine. The sun produces solar energy through a process called the thermonuclear process in the form of heat and electromagnetic radiations. These electromagnetic radiations contain the energy that reaches the earth. So, solar energy is a type of indirect source of energy so to utilize it completely we require two components: first, the collector which can collect radiations that are coming from the sun and transform them into electrical energy, second, a storage unit as radiations are differing in nature. But during the implementation of the solar power plants on land, complications are faced by the government and associates of the project such as land availability, land development, and land procurement, timely evacuation of the land, and many more. To conquer these difficulties an ingenious idea has come forward for setting up solar power plants on the water that is on ponds, irrigation canal tops, water bodies like reservoirs, quarry lakes, dam backwater, which are generally owned by the government. And this innovative idea is known as Floating Solar Power Plants.

II. FUNDAMENTALS OF FLOATING SOLAR POWER PLANTS

Floating solar or floating photovoltaics (FPV), also called floatovoltaics are actually the solar panels which are installed on floating platforms like rafts and anchored firmly to avoid the damages caused by even the worse weather conditions. The solar panels are installed on floating platforms like rafts that are anchored firmly to avoid the damages caused by even the worse weather conditions. Furthermore, research suggests that solar panels that are installed on land surfaces cause minimization of yields, as the ground gets heated up and then this affects the rear surfaces of solar panels. On the other side, if the rear surfaces of solar panels are placed on top of a water body, the solar panels will be able to cool themselves more efficiently as a result they last longer. Also, floating solar panels shade the water over which they float which reduces the evaporation rate by up to

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:10/October-2021 Impact Factor- 6.752 www.irjmets.com

technology with a floating mechanism result in photovoltaic (PV) floating power generation. This combination of a new concept consists of a floating system(structure + floater) that enables the installation of the PV module, PV system similar to electrical junction boxes, that are installed on top of it along with an underwater cable which transfers the generated power to the PV system development.

Fig-1: Floating Solar Power Plant

A small number of such systems have been installed in Singapore, India, Japan, France, the United States, South Korea and the United Kingdom. Unlike most solar plants installed on land, floating arrays can be unnoticeable because they get hidden from public view. They achieve higher productivity than PV panels installed on land because water cools the panels. These panels also have a special coating to protect them from rust or corrosion.

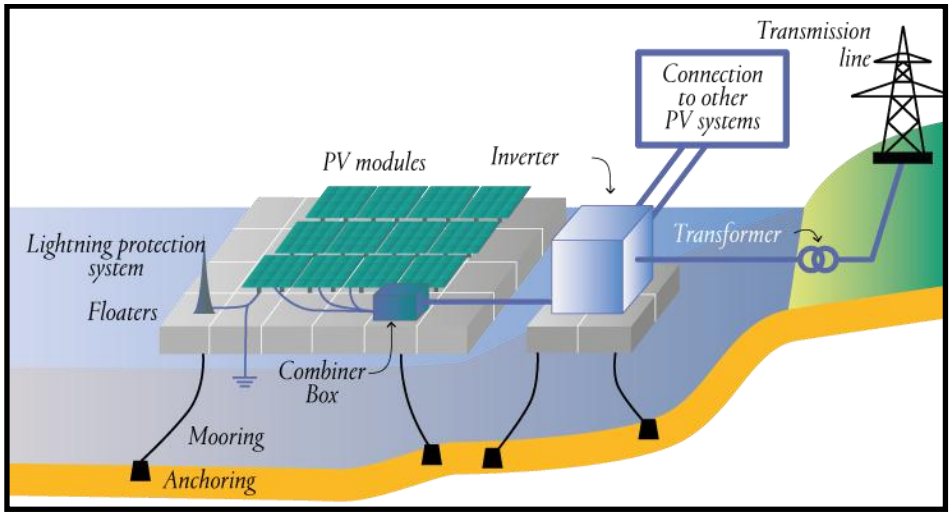


Fig-2: Floating power plant outline

III. COMPONENTS OF FLOATING SOLAR POWER PLANTS

The floating Photovoltaic system is a recently developed technique of generating solar energy installed on water surfaces available on dams, reservoirs, and other water bodies. The floating solar panel consists of a floating system, mooring system, PV system, and underwater cables.

1. **Pontoon/Floating Structure**: It is a floating structure. Pontoon has buoyancy enough to float on water and hold up a heavy load. These structures allow installation of PV module on them.
2. **Mooring Structure:** It is a permanent body that secures the floating structure. The mooring structure prevents the free movement of floating structure over the water. An anchor mooring fixes a floating structure's location relative to a point on the bottom of a canal without connecting the floating structure to shore.
3. **Solar Module**: Solar module is PV Generation equipment(similar to electric junction boxes) which are installed on top of the floating system. A single solar module can produce only a small amount of power therefore most of the installations comprise several modules. A PV system typically contains an array of

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:10/October-2021 Impact Factor- 6.752 www.irjmets.com

solar modules, a solar inverter, a battery and/or solar tracker, and interconnection wiring. Generally, crystalline solar PV modules are used for floating solar systems.

4. **Cabling:** Cabling enables the transfer of generated power from water bodies to the substation. Due to their outdoor usage, these solar cables are specially designed to be resistant against UV radiation and extremely high temperature fluctuations

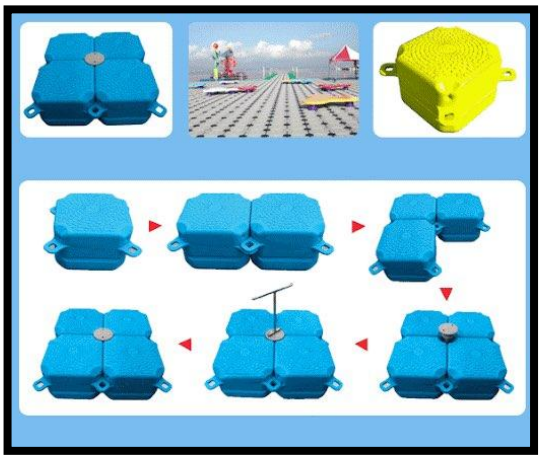


Fig-3: Pontoon structure

IV. MERITS OF FLOATING SOLAR POWER PLANTS

1. Due to the cooling effect of the water, floating solar power plant system in most cases produces more electricity than ground-mount and rooftop systems.
2. They are designed and engineered in such a way that they can withstand extreme physical stress, including cyclone and storm conditions.
3. Land is not used in this case and so there is no evacuation of people and animals too, which conserves the local environment.
4. Geographically any water bodies with plentiful sunlight can be used to install floating power plants.
5. These installations minimize the water evaporation rate and algae growth as they shade the water.
6. These floating platforms are 100% recyclable as they utilize high-density polyethylene, which can withstand ultraviolet rays and resist corrosion.
7. Easy to build and faster deployment.

V. DEMERITS OF FLOATING SOLAR POWER PLANTS

1. It is a relatively developing and immature technology.
2. Lack of knowledge and experience in handling this technology.
3. Solar panels get damaged due to high waves and saltwater overtime
4. Long-term maintenance requirements and durability of floating solar PV are yet to be seen.
5. Lack of cooperation from local distribution utility.

VI. FUTURE SCOPE OF FLOATING POWER PLANT

1. This technology is expanding and developing in countries like India, China, Thailand, Malaysia and other developing countries due to emergence of new markets and investment.
2. Availability of trained manpower and Government Policies has uplifted the conviction of investors.
3. Prevailing land issues and availability of water bodies has prompted investors to shift towards floating solar power plants.
4. Stable floating PV platforms results in minimum operation & maintenance cost.
5. Growing innovations in Floating technology .

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:03/Issue:10/October-2021 Impact Factor- 6.752 www.irjmets.com

VII. ENVIRONMENTAL EFFECTS OF FLOATING POWER PLANT

A Floating solar platform enables standard Photovoltaic panels to be installed on large water bodies such as dam backwater, reservoirs, lakes, irrigation canals, or remediation and ponds. An uncomplicated and inexpensive floating solar platform is especially well suitable for energy and water-intensive industries that cannot afford to waste either land or water. Wineries, dairy farms, fish farms, mining companies, wastewater treatment plants, irrigation districts, and water agencies are industries that can benefit from the synergy that floating solar system creates between water and sun.

VIII. CONCLUSION

The review presented in this paper shows the course of events of ideas and drifting sun-oriented PV projects that have been established to date. These frameworks were either built for re-search purposes or for business use. All lattice associated frameworks are kept above water utilizing boats or floats with boards unbendingly associated with these floats. Floating solar application with challenges and opportunities has been discussed.

- The examination introduced in this paper can be used as instrument for future advancement of drifting photovoltaic frameworks.
- To change drifting sun based, dangers distinguished should be followed properly. Be that as it may, the future appears to be splendid for the drifting sunlight-based innovation.
- Drifting solar technology would end up being an imaginative advance as it could take care of the enduring issue of land.
- In India huge water bodies are accessible in Eastern, Southern and South-eastern piece of the country in states, for example, West Bengal, Assam, Orissa and Andhra Pradesh, Tamil Nadu and Kerala. This innovation can be embraced in these states prompting extensive investment funds ashore costs and cut down power age costs, hence diminishing the hole among warm and sun-oriented power.
- The impact of salt water on the PV structure and the module execution must be investigated.
- Improvements of huge megawatt scale drifting sun powered ranches in not-so-distant future might pay way for the seaward sun-based innovation improvement.
- A large portion of the tasks in presence in corporate inflexible glasslike PV modules which are unequipped for withstanding cruel water climate along these lines research on adaptable slender film innovation for such unforgiving condition must be investigated.
- The proficiency of drifting sun-based plant is 11% higher and lessens the water vanishing by 70%, but the venture of such power plant is 1.2% occasions higher than the traditional sunlight-based power plant.
- Greatest speed of wind, water ebb and flow, temperature limit, snow burden, twister and tropical storm must be thought of while planning the sunlight powered charger.
- Fitting wellbeing measures to ship the power from the water bodies to the land must be completed.

IX. REFERENCES

[1] Hyder Farhan, Sudhakar K, Rizalman Mamat. Solar PV tree design: A review. Renewable and Sustainable Energy Reviews. 2018;82(2018):1079–1096.

[2] Deepak Bishoyi, Sudhakar K. Modeling and performance simulation of 100 MW LFR based solar thermal power plant in Udaipur India. Resource-Efficient Technologies. 2017;10:216–226.

[3] Akash Kumar Shukla, Sudhakar K, Prashant Baredar, et al. BIPV in Southeast Asian countries – opportunities and challenges. Renewable Energy Focus. 21;25–32.

[4] Armstrong, Robert (12 November 2014). "The Case for Solar Energy Parking Lots". Absolute Steel. Retrieved 15 November 2014.

[5] Sunflower Floating Solar Power Plant In Korea". CleanTechnica. 21 December 2014.

[6] Submerging Solar: New Frontiers for Renewable Energy (Online) http://www.ecomagination.com/submerging-solar-new-frontiers-for-renewable- energy

[7] Jinyoung Song and Yosoon Choi, "Analysis of the Potential for Use of Floating Photovoltaic Systems on Mine Pit Lakes: Case Study at the Ssangyong Open-Pit Limestone Mine in Korea", Energies 2016, 9, 102; doi:10.3390/en9020102.

[8] "Sunflower Floating Solar Power Plant in Korea". CleanTechnica. 21 December 2014 "Napa Winery Pioneers Solar Floatovoltaics". Forbes. 18 April 2012. Re- trieved 31 May 2013.

[9] G.M. Tina, M. Rosa-Clot, P. Rosa-Clot, "Electrical behaviour and optimization of panels and reflector of a photovoltaic floating plant", Proceedings of the 26th Euro- pean Photovoltaic Solar Energy Conference and Exhibition (EU PVSEC'11) 2011. 4371–4375.

[10] Y. K. Choi, N. H. Lee and K. J. Kim, "Empirical Research on the efficiency of Floating PV systems compared with Overland PV Systems", CES-CUBE 2013, Guam, USA, (2013) July 19.

arXiv:2112.07073v1 [math.CV] 14 Dec 2021

ANGULAR ESTIMATES FOR CERTAIN ANALYTIC FUNCTIONS

POOJA YADAV AND S. SIVAPRASAD KUMAR

Abstract. For $\alpha, \beta \in (-1, 1]$, let $\mathcal{G}(\alpha, \beta)$ be the class of all analytic functions h, satisfying $-\beta\pi/2 < \arg h(z) < \alpha\pi/2$ for all $|z| < 1$. Using Nunokawa's Lemma, we obtain sufficient conditions for functions to be in $\mathcal{G}(\alpha, \beta)$, which involves slit domains. Several special cases of our results are also pointed out here.

1. Introduction

Let $\mathbb{D}$ be the open unit disk in the complex plane. For $n \in \mathbb{N}$ and $a \in \mathbb{C}$, let $\mathcal{H}[a, n]$ denote the class of functions f analytic in $\mathbb{D}$ with $f(z) = a + a_n z^n + a_{n+1} z^{n+1} + \cdots$. Let $\mathcal{H}_1 := \mathcal{H}[1, 1]$ and $\mathcal{A}$ be the subclass of $\mathcal{H}[0, 1]$ consisting of functions normalized by $f'(0) = 1$. A subclass of $\mathcal{A}$ consisting of all univalent functions is denoted by $\mathcal{S}$ and further $\mathcal{S}^* \subset \mathcal{S}$ consists of functions f that are starlike with respect to origin. A function which is starlike with respect to each of its points is known as a convex function and the set of all convex functions is denoted by $\mathcal{C}$. Analytically $f \in \mathcal{S}^*$ ($f \in \mathcal{C}$) if, and only if, $f \in \mathcal{S}$ and $\mathrm{Re}(zf'(z)/f(z)) > 0$ ($\mathrm{Re}(zf''(z)/f'(z)) > -1$). For two analytic functions f and g, f is subordinate to g, written as $f \prec g$, if there exists a Schwarz function $w(z)$ such that $f(z) = g(w(z))$. For $-\pi/2 < \lambda < \pi/2$, the tilted Carathéodory class by angle λ is defined as:

$$\mathcal{P}_\lambda := \{p \in \mathcal{H}_1 : \mathrm{Re}(e^{i\lambda} p(z)) > 0\}.$$

It is well known that $\mathcal{P}_0 =: \mathcal{P}$, the Carathéodory class, which consists of $p \in \mathcal{H}_1$ such that $p(z) \prec (1 + z)/(1 - z)$. In 1973 Janowski[4], introduced and studied the class $\mathcal{P}(A, B)$ with $-1 \le B < A \le 1$, which is the general form of the class $\mathcal{P}$ consisting of functions $p \in \mathcal{P}$ such that $p(z) \prec (1 + Az)/(1 + Bz)$. Later many researches have studied and generalised this class in various ways, for instance, see [8, 9, 2, 10]. $\mathcal{SS}^*(\alpha)$ for $\alpha \in (0, 1]$ is the class of strongly starlike functions of order α, which is defined as:

$$\mathcal{SS}^*(\alpha) = \left\{ f \in \mathcal{S} : \left| \arg \frac{zf'(z)}{f(z)} \right| < \frac{\alpha\pi}{2}, \text{ equivalently } \frac{zf'(z)}{f(z)} \prec \left(\frac{1+z}{1-z} \right)^\alpha \right\}.$$

Definition 1.1. Let $\alpha, \beta \in (-1, 1]$ and a function $h(z) = \sum_{n=0}^{\infty} h_n z^n$ analytic in $\mathbb{D}$ is said to be in the class $\mathcal{G}(\alpha, \beta)$, if

$$-\frac{\beta\pi}{2} < \arg h(z) < \frac{\alpha\pi}{2}. \tag{1.1}$$

In the present investigation, we obtain some sufficient conditions involving slit domains, for functions to be in the class $\mathcal{G}(\alpha, \beta)$. Further special cases of our results also discussed here.

2. Preliminary

We list below a few results that are required to prove our main results. The following lemma is a restatement of Nunokawa's Lemma, given in [5, 11].

2010 *Mathematics Subject Classification.* 30C45,30C50, 30C80.

Key words and phrases. Analytic function, Argument , Subordination , Slit Domain.

Lemma 2.1. [11] *Let $h \in \mathcal{H}[1,n]$. If there exist two points $z_1, z_2 \in \mathbb{D}$ such that*

$$-\frac{\beta\pi}{2} = \arg h(z_1) < \arg h(z) < \arg h(z_2) = \frac{\alpha\pi}{2}$$

for $\alpha, \beta \in (-1,1]$ and $|z| < |z_1| = |z_2|$. Then

$$\frac{z_j h'(z_j)}{h(z_j)} = (-1)^j i \frac{\alpha+\beta}{2} k_j \quad (j=1,2),$$

where

$$\left(\frac{k_j}{n} + (-1)^{j+1}\tan\eta\right)\cos\eta \geq \left(x_j + \frac{1}{x_j}\right)\frac{1}{2} \quad (x_j > 0),$$

$$\exp(-i\eta)(h(z_j))^{\frac{2}{\alpha+\beta}} = (-1)^j i x_j \text{ and } \eta = \frac{(\alpha-\beta)\pi}{(\alpha+\beta)2}.$$

Lemma 2.2. [2] *For $0 < \alpha \leq 1$ and $-1 < m < 1$, then the function*

$$h(z) = \left(\frac{1+e^{im\pi}z}{1-z}\right)^{\alpha},$$

is analytic, univalent and convex in $\mathbb{D}$ with

$$h(\mathbb{D}) = \left\{w \in \mathbb{C} : -\alpha(1-m)\frac{\pi}{2} \leq \arg w \leq \alpha(1+m)\frac{\pi}{2}\right\}.$$

Lemma 2.3. [2] *Let $h \in \mathcal{H}_1$ and $0 \leq b \leq 1$ with $b + e^{im\pi} \neq 0$, where $-1 \leq m \leq 1$. Also, if*

$$h(z) \prec \frac{1+e^{im\pi}z}{1-bz},$$

then

$$\mathrm{Re}(e^{-i\lambda}h(z)) > 0,$$

where $\lambda = \tan^{-1}\left(\dfrac{b\sin(m\pi)}{b\cos(m\pi)+1}\right)$.

3. Main Results

We begin with the following theorem:

Theorem 3.1. *Let $\alpha, \beta \in (-1,1]$ and $h \in \mathcal{H}[1,n]$. Then $h \in \mathcal{G}(\alpha,\beta)$, whenever*

$$(h(z))^{\frac{2}{\alpha+\beta}} + \frac{zh'(z)}{h(z)} \neq \{x_1 + iy_1\} \cup \{x_2 + iy_2\}, \tag{3.1}$$

where $|z| < 1$,

$$y_1 \leq -\frac{(\alpha+\beta)n}{2\cos\eta}\left(\sqrt{1+\frac{4\cos^2\eta}{(\alpha+\beta)n}} - \sin\eta\right),$$

$$y_2 \geq \frac{(\alpha+\beta)n}{2\cos\eta}\left(\sqrt{1+\frac{4\cos^2\eta}{(\alpha+\beta)n}} + \sin\eta\right),$$

$$x_j = (-1)^{j+1}\frac{\sin\eta}{\sqrt{1+\frac{4\cos^2\eta}{(\alpha+\beta)n}}} \quad (j=1,2) \quad \text{and} \quad \eta = \left(\frac{\alpha-\beta}{\alpha+\beta}\right)\frac{\pi}{2}.$$

Proof. If there exist z_1 and z_2 with $|z_1|=|z_2|=r<1$, such that

$$-\frac{\beta\pi}{2}=\arg h(z_1)<\arg h(z)<\arg h(z_2)=\frac{\alpha\pi}{2},$$

for $|z|<r$. Then by Lemma 2.1, we get

$$\frac{z_jh'(z_j)}{h(z_j)}=(-1)^ji\frac{\alpha+\beta}{2}k_j\quad(j=1,2),$$

where for $x_j>0$, we have

$$\frac{k_j}{n}\geq\left(x_j+\frac{1}{x_j}\right)\frac{\sec\eta}{2}+(-1)^j\tan\eta\ \text{ and }\ \exp(-i\eta)(h(z_j))^{\frac{2}{\alpha+\beta}}=(-1)^jix_j.$$

For $z=z_1$, we have

$$\operatorname{Re}\left(h(z_1)^{\frac{2}{\alpha+\beta}}+\frac{z_1h'(z_1)}{h(z_1)}\right)=x_1\sin\eta$$

and

$$\begin{aligned}\operatorname{Im}\left(h(z_1)^{\frac{2}{\alpha+\beta}}+\frac{z_1h'(z_1)}{h(z_1)}\right)&=-x_1\cos\eta-\frac{\alpha+\beta}{2}k_1\\&\leq-x_1\cos\eta-\frac{\alpha+\beta}{2}n\left(\frac{1}{2}\left(x_1+\frac{1}{x_1}\right)\sec\eta-\tan\eta\right)\\&=-n\frac{\alpha+\beta}{2}\cos\eta f(x_1),\end{aligned}$$

where

$$f(x):=\frac{2x}{(\alpha+\beta)n}+\left(x+\frac{1}{x}\right)\frac{1}{2}\sec^2\eta-\frac{\sin\eta}{\cos^2\eta}\quad(x>0).$$

By a calculation, we see that $f(x)$ attains its minimum at $x:=x_1=\left(1+\dfrac{4\cos^2\eta}{n(\alpha+\beta)}\right)^{-\frac{1}{2}}$. Thus

$$\operatorname{Im}\left(h(z_1)^{\frac{2}{\alpha+\beta}}+\frac{z_1h'(z_1)}{h(z_1)}\right)\leq-\frac{(\alpha+\beta)n}{2\cos\eta}\left(\sqrt{1+\frac{4\cos^2\eta}{(\alpha+\beta)n}}-\sin\eta\right)$$

and

$$\operatorname{Re}\left(h(z_1)^{\frac{2}{\alpha+\beta}}+\frac{z_1h'(z_1)}{h(z_1)}\right)=\frac{\sin\eta}{\sqrt{1+\frac{4\cos^2\eta}{n(\alpha+\beta)}}},$$

which contradicts (3.1). After going through the same methodology for $z=z_2$, we obtain

$$\operatorname{Im}\left(h(z_2)^{\frac{2}{\alpha+\beta}}+\frac{z_2h'(z_2)}{h(z_2)}\right)\geq\frac{(\alpha+\beta)n}{2\cos\eta}\left(\sqrt{1+\frac{4\cos^2\eta}{(\alpha+\beta)n}}+\sin\eta\right)$$

and

$$\operatorname{Re}\left(h(z_2)^{\frac{2}{\alpha+\beta}}+\frac{z_2h'(z_2)}{h(z_2)}\right)=-\frac{\sin\eta}{\sqrt{1+\frac{4\cos^2\eta}{n(\alpha+\beta)}}},$$

again a contradiction, which eventually proves the result. ∎

Remark 3.2. If we take $\alpha=\beta\in(0,1]$ in Theorem 3.1, it reduces to a result of Nunokawa and Sokół [7, Theorem 2.1].

Further, we obtain the following result form Theorem 3.1:

Corollary 3.3. *Let $\lambda \in [0,1)$ and $f \in \mathcal{A}$. If*

$$\lambda \frac{zf'(z)}{f(z)} + (1-\lambda)\left(1 + \frac{zf''(z)}{f'(z)}\right) =: R_\lambda(z)$$

lies in the complex plane with slits along $\operatorname{Re} w = 0$ *and* $|\operatorname{Im} w| \geq (1-\lambda)(\sqrt{1+2/(1-\lambda)}) =: C(\lambda)$. *Then we have*

$$\operatorname{Re} \frac{zf'(z)}{f(z)} > 0, \quad \forall z \in \mathbb{D}.$$

Proof. Letting $h(z) = (zf'(z)/f(z))^{1-\lambda}$, $\alpha = \beta = 1-\lambda$ and $n = 1$ in Theorem 3.1, we have

$$(h(z))^{\frac{2}{\alpha+\beta}} + \frac{zh'(z)}{h(z)} = \lambda \frac{zf'(z)}{f(z)} + (1-\lambda)\left(1 + \frac{zf''(z)}{f'(z)}\right) \neq 0 + iy,$$

where y is such that $|y| \geq (1-\lambda)(\sqrt{1+2/(1-\lambda)})$. Hence the result follows from Theorem 3.1. ∎

Note that, the corollary 3.3 with $\lambda = 0$ reduces to the following special case of Open Door Lemma (see [3]).

Corollary 3.4. *Let $f \in \mathcal{A}$ and $1 + \frac{zf''(z)}{f'(z)}$ lies in the complex plane with slits along* $\operatorname{Re} w = 0$ *and* $|\operatorname{Im} w| \geq \sqrt{3}$. *Then*

$$\operatorname{Re} \frac{zf'(z)}{f(z)} > 0, \quad \forall z \in \mathbb{D}.$$

Remark 3.5. If $\lambda_1, \lambda_2 \in [0,1)$ with $\lambda_1 \leq \lambda_2$, we have $C(\lambda_1) \geq C(\lambda_2)$ and $C(0) = \sqrt{3}$. This leads to $R_{\lambda_2}(z) \prec R_{\lambda_1}(z) \prec R_0(z)$.

Theorem 3.6. *Let $\lambda \in [0, \pi/2)$ and $h \in \mathcal{H}_1$. Then $h \in \mathcal{P}_{-\lambda}$, whenever for $|z| < 1$ and*

$$\forall A \in \mathbb{R} \quad : \quad |A| \geq \sec\lambda \sqrt{1 + 2\cos\lambda} - \tan\lambda,$$

we have

$$e^{-i\lambda} h(z) + \frac{zh'(z)}{h(z)} \neq iA. \tag{3.2}$$

Proof. If $h \in \mathcal{P}_{-\lambda}$ doesn't hold then by Lemma 2.2 and Lemma 2.3, we have $h(z) \not\prec \frac{1+e^{i2\lambda}z}{1-z}$. Therefore there exist z_1 and z_2 with $|z_1| = |z_2| = r < 1$, such that

$$-\left(\frac{\pi}{2} - \lambda\right) = \arg h(z_1) < \arg h(z) < \arg h(z_2) = \frac{\pi}{2} + \lambda,$$

for $|z| < r$. Now by Lemma 2.1, we get

$$\frac{z_j h'(z_j)}{h(z_j)} = (-1)^j i k_j \quad (j = 1, 2),$$

where

$$k_j \geq \frac{1}{2}\left(x_j + \frac{1}{x_j}\right)\sec\lambda + (-1)^j \tan\lambda \quad (x_j > 0)$$

and

$$e^{-i\lambda} h(z_j) = (-1)^j i x_j \quad (j = 1, 2).$$

Now for $z = z_1$, we have

$$e^{-i\lambda} h(z_1) + \frac{z_1 h'(z_1)}{h(z_1)} = -i(x_1 + k_1).$$

Here

$$\operatorname{Re}\left(e^{-i\lambda}h(z_1)+\frac{z_1h'(z_1)}{h(z_1)}\right)=0 \tag{3.3}$$

and

$$\operatorname{Im}\left(e^{-i\lambda}h(z_1)+\frac{z_1h'(z_1)}{h(z_1)}\right)\leq-\left(x_1+\frac{1}{2}\left(x_1+\frac{1}{x_1}\right)\sec\lambda-\tan\lambda\right)=:-f(x_1).$$

By a calculation we obtain that $f(x)$ $(x>0)$ attains its minimum at $x=(1+2\cos\lambda)^{-1/2}$. Thus

$$\operatorname{Im}\left(e^{-i\lambda}h(z_1)+\frac{z_1h'(z_1)}{h(z_1)}\right)\leq-\left(\sec\lambda\sqrt{1+2\cos\lambda}-\tan\lambda\right). \tag{3.4}$$

Clearly, (3.3) and (3.4) contradicts the hypothesis given in (3.2). Similarly for $z=z_2$, we have

$$e^{-i\lambda}h(z_2)+\frac{z_2h'(z_2)}{h(z_2)}=i(x_2+k_2)$$

Here

$$\operatorname{Re}\left(e^{-i\lambda}h(z_2)+\frac{z_2h'(z_2)}{h(z_2)}\right)=0 \tag{3.5}$$

and

$$\operatorname{Im}\left(e^{-i\lambda}h(z_2)+\frac{z_2h'(z_2)}{h(z_2)}\right)\geq x_2+\frac{1}{2}\left(x_2+\frac{1}{x_2}\right)\sec\lambda+\tan\lambda=:g(x_2).$$

By a calculation we obtain that $g(x)$ $(x>0)$ attains its minimum at $x=(1+2\cos\lambda)^{-1/2}$. Thus

$$\begin{aligned}\operatorname{Im}\left(e^{-i\lambda}h(z_2)+\frac{z_2h'(z_2)}{h(z_2)}\right)&\geq\sec\lambda\sqrt{1+2\cos\lambda}+\tan\lambda\\&\geq\sec\lambda\sqrt{1+2\cos\lambda}-\tan\lambda.\end{aligned} \tag{3.6}$$

Again (3.5) and (3.6) contradicts the hypothesis given in (3.2). Hence the result. ∎

Now the following are some corollaries of Theorem 3.6:

Corollary 3.7. *Let $\alpha\in[0,1)$, $\lambda\in[0,\pi/2)$ and $p(z)\in\mathcal{H}_1$. If*

$$\frac{p(z)-\alpha}{1-\alpha}+e^{i\lambda}\frac{zp'(z)}{p(z)-\alpha}\prec e^{i\lambda}\frac{2Az}{1-z^2},$$

where A is as given in Theorem 3.6, then

$$\operatorname{Re}(e^{-i\lambda}p(z))>\alpha\cos\lambda.$$

Proof. Taking $h(z)=\dfrac{p(z)-\alpha}{1-\alpha}$ in (3.2), we obtain

$$e^{-i\lambda}h(z)+\frac{zh'(z)}{h(z)}=e^{-i\lambda}\frac{p(z)-\alpha}{1-\alpha}+\frac{zp'(z)}{p(z)-\alpha}\neq iA,$$

which can also be restated as

$$\frac{p(z)-\alpha}{1-\alpha}+e^{i\lambda}\frac{zp'(z)}{p(z)-\alpha}\prec e^{i\lambda}\frac{2Az}{1-z^2}.$$

Hence the result follows from Theorem 3.6. ∎

Remark 3.8. For $f(z)\in\mathcal{A}_1$, letting $h(z)=zf'(z)/f(z)$ and $f'(z)$ respectively in Theorem 3.6, we obtain, if

(*i*) $1+\dfrac{zf''(z)}{f'(z)}+(e^{-i\lambda}-1)\dfrac{zf'(z)}{f(z)}\prec\dfrac{2Az}{1-z^2}$, then $\operatorname{Re}\left(e^{-i\lambda}\dfrac{zf'(z)}{f(z)}\right)>0$.

(*ii*) $\operatorname{Re}\left(e^{-i\lambda}\dfrac{zf'(z)}{f(z)}\right)>0$. then $f'(z)\prec e^{i\lambda}\dfrac{1+z}{1-z}$.

Our next corollary is a collection of many results obtained by taking $\lambda = 0$ and $h(z) = zf'(z)/f(z)$, $f'(z)$, $zf'(z)/g(z)$ and $zf'(z)/(f^{(1-\beta)}(z)g^{\beta}(z))$ respectively in Theorem 3.6.

Corollary 3.9. *Let $f, g \in \mathcal{A}_1$ with $g \in \mathcal{S}^*$ and $\beta \in [0,1)$. If*

(i) $1 + \dfrac{zf''(z)}{f'(z)} \prec \dfrac{2\sqrt{3}z}{1-z^2}$, *then* $\operatorname{Re} \dfrac{zf'(z)}{f(z)} > 0$ *(special case of Open Door Lemma (see [3]))*.

(ii) $f'(z) + \dfrac{zf''(z)}{f'(z)} \prec \dfrac{2\sqrt{3}z}{1-z^2}$, *then* $\operatorname{Re} f'(z) > 0$.

(iii) $\dfrac{zf''(z)}{f'(z)} - \dfrac{zg'(z)}{g(z)} + \dfrac{zf'(z)}{g(z)} \prec \dfrac{2\sqrt{3}z}{1-z^2}$, *then* $\operatorname{Re} \dfrac{zf'(z)}{g(z)} > 0$.

(iv) $1 + \dfrac{zf''(z)}{f'(z)} - (1-\beta)\dfrac{zf'(z)}{f(z)} - \beta\dfrac{zg'(z)}{g(z)} + \dfrac{zf'(z)}{f^{(1-\beta)}(z)g^{\beta}(z)} \prec \dfrac{2\sqrt{3}z}{1-z^2}$, *then* $\operatorname{Re} \dfrac{zf'(z)}{f^{(1-\beta)}(z)g^{\beta}(z)} > 0$.

Proof. Taking $\lambda = 0$ and $h(z) = zf'(z)/f(z)$ in (3.2), we obtain

$$h(z) + \frac{zh'(z)}{h(z)} = 1 + \frac{zf''(z)}{f'(z)} \neq i\sqrt{3}.$$

Hence by Theorem 3.6 we obtain (i). Similarly, we have (ii), (iii) and (iv). ∎

In [1, Theorem 3.1], authors gave some sufficient condition for an analytic function $h \in \mathcal{H}_1$ to have argument bounds as $|\arg h(z)| < \beta\pi/2$, where $\beta \in (0,1)$. Here in the following theorem, we obtain sufficient condition for $h \in \mathcal{G}(\alpha, \beta)$.

Theorem 3.10. *Let $\alpha, \beta \in (-1,1)$ and $0 < \gamma \leq 1$. If $h \in \mathcal{H}_1$ then $h \in \mathcal{G}(\alpha,\beta)$, whenever*

$$\delta_1 \frac{\pi}{2} < \arg h(z) + \gamma \arg\left(1 + \frac{zh'(z)}{h^2(z)}\right) < \delta_2 \frac{\pi}{2}, \tag{3.7}$$

where

$$\delta_1 = -\left(\beta + \frac{2\gamma}{\pi} \tan^{-1} \frac{(\alpha+\beta)\sin\frac{(1-\beta)\pi}{2}}{(\alpha+\beta)\cos\frac{(1-\beta)\pi}{2} + M_1(\alpha,\beta)}\right),$$

$$\delta_2 = \alpha + \frac{2\gamma}{\pi} \tan^{-1} \frac{(\alpha+\beta)\sin\frac{(1-\alpha)\pi}{2}}{(\alpha+\beta)\cos\frac{(1-\alpha)\pi}{2} + M_2(\alpha,\beta)},$$

$$M_j(\alpha,\beta) = N\left(\frac{(-1)^j(\alpha+\beta)}{2-\alpha-\beta}\sin\eta + \sqrt{\frac{2+\alpha+\beta}{2-\alpha-\beta}}\cos\eta\right) \quad (j = 1,2),$$

$$N(x) = 4x^{\frac{\alpha+\beta}{2}}\left(\left(x + \frac{1}{x}\right)\sec\eta + (-1)^j 2\tan\eta\right)^{-1} \text{ and } \eta = \frac{(\alpha-\beta)\pi}{(\alpha+\beta)2}.$$

Proof. If there exist z_1 and z_2 with $|z_1| = |z_2| = r < 1$, such that

$$-\frac{\beta\pi}{2} = \arg h(z_1) < \arg h(z) < \arg h(z_2) = \frac{\alpha\pi}{2},$$

for $|z| < r$. Then by Lemma 2.1, we get

$$\frac{z_j h'(z_j)}{h(z_j)} = (-1)^j i \frac{\alpha+\beta}{2} k_j, \quad (j = 1,2),$$

where for $x_j > 0$

$$k_j \geq \left(\frac{1}{x_j} + x_j\right)\frac{\sec\eta}{2} + (-1)^j \tan\eta \text{ and } h(z_j) = \left((-1)^j i x_j \exp(i\eta)\right)^{\frac{\alpha+\beta}{2}}.$$

For $z = z_1$, we have

$$h(z_1)\left(1+\frac{z_1h'(z_1)}{h^2(z_1)}\right)^{\gamma} = x^{\frac{\alpha+\beta}{2}}e^{-i\frac{\beta\pi}{2}}\left(1+\frac{e^{-i(1-\beta)\frac{\pi}{2}}(\alpha+\beta)k_1}{2x^{\frac{\alpha+\beta}{2}}}\right)^{\gamma}.$$

Moreover,

$$\begin{aligned}\arg h(z_1)+\gamma\arg\left(1+\frac{z_1h'(z_1)}{h^2(z_1)}\right) &= -\frac{\beta\pi}{2}+\gamma\tan^{-1}-\frac{\sin\left((1-\beta)\frac{\pi}{2}\right)(\alpha+\beta)k_1}{2x_1^{\frac{\alpha+\beta}{2}}+\cos\left((1-\beta)\frac{\pi}{2}\right)(\alpha+\beta)k_1},\\ &\leq -\frac{\beta\pi}{2}-\gamma\tan^{-1}\frac{\sin\left((1-\beta)\frac{\pi}{2}\right)(\alpha+\beta)}{N(x_1)+\cos\left((1-\beta)\frac{\pi}{2}\right)(\alpha+\beta)},\end{aligned}$$

where $N(x)$ is defined in hypothesis. As $x > 0$, by a calculation, we see that $N(x)$ attains its maximum at

$$x := x_1 = \frac{-(\alpha+\beta)}{2-\alpha-\beta}\sin\eta+\sqrt{\frac{2+\alpha+\beta}{2-\alpha-\beta}}\cos\eta.$$

Thus

$$\arg h(z_1)+\gamma\arg\left(1+\frac{z_1h'(z_1)}{h^2(z_1)}\right)\leq -\frac{\beta\pi}{2}-\gamma\tan^{-1}\frac{\sin\left((1-\beta)\frac{\pi}{2}\right)(\alpha+\beta)}{M_1(\alpha,\beta)+\cos\left((1-\beta)\frac{\pi}{2}\right)(\alpha+\beta)},$$

which contradicts (3.7). Similarly for $z = z_2$, we have

$$\arg h(z_2)+\gamma\arg\left(1+\frac{z_2h'(z_2)}{h^2(z_2)}\right)\geq \frac{\alpha\pi}{2}+\gamma\tan^{-1}\frac{\sin\left((1-\alpha)\frac{\pi}{2}\right)(\alpha+\beta)}{M_2(\alpha,\beta)+\cos\left((1-\alpha)\frac{\pi}{2}\right)(\alpha+\beta)},$$

which again contradicts (3.7). Thus result holds. ∎

Further, we have the following interesting corollaries from Theorem 3.10:

Corollary 3.11. *Let $\alpha,\beta\in(-1,1)$ and $0<\gamma\leq 1$. Moreover, $f\in\mathcal{A}$, which satisfies*

$$\frac{\delta_1\pi}{2} < (1-\gamma)\arg\left(\frac{zf'(z)}{f(z)}\right)+\gamma\arg\left(1+\frac{zf''(z)}{f'(z)}\right) < \frac{\delta_2\pi}{2}, \tag{3.8}$$

where δ_1 and δ_2 are defined in Theorem 3.10. Then

$$-\frac{\beta\pi}{2} < \arg\frac{zf'(z)}{f(z)} < \frac{\alpha\pi}{2},$$

or can say

$$\frac{zf'(z)}{f(z)} \prec \left(\frac{1+e^{i(\frac{\alpha-\beta}{\alpha+\beta})\pi}z}{1-z}\right)^{(\alpha+\beta)/2}.$$

Proof. Taking $h(z) = zf'(z)/f(z)$ in Theorem 3.10, we obtain

$$\left(\frac{zf'(z)}{f(z)}\right)^{1-\gamma}\left(1+\frac{zf''(z)}{f'(z)}\right)^{\gamma} = h(z)\left(1+\frac{zh'(z)}{h^2(z)}\right)^{\gamma}.$$

Now by (3.7) and (3.8), we have

$$\frac{\delta_1\pi}{2} < (1-\gamma)\arg\left(\frac{zf'(z)}{f(z)}\right)+\gamma\arg\left(1+\frac{zf''(z)}{f'(z)}\right) = \arg h(z)+\gamma\arg\left(1+\frac{zh'(z)}{h^2(z)}\right) < \frac{\delta_2\pi}{2}.$$

Thus the result follows at once from Theorem 3.10. ∎

By taking $h(z) = zf'(z)/f(z)$ and $\alpha = \beta$ in Theorem 3.10, we obtain the following result of Nunokawa and Sokół[6] with an alternate proof:

Corollary 3.12. *Let $f \in \mathcal{A}$, $\alpha \in (0,1)$ and $0 < \gamma \leq 1$. Moreover, f satisfies*

$$\left| \arg \left\{ \left(\frac{zf'(z)}{f(z)} \right)^{1-\gamma} \left(1 + \frac{zf''(z)}{f'(z)} \right)^{\gamma} \right\} \right| < \frac{\delta\pi}{2},$$

where

$$\delta = \alpha + \frac{2\gamma}{\pi} \tan^{-1} \frac{2\alpha \sin \frac{(1-\alpha)\pi}{2}}{2\alpha \cos \frac{(1-\alpha)\pi}{2} + M(\alpha)},$$

with

$$M(\alpha) = \frac{4}{(\frac{1+\alpha}{1-\alpha})^{(1-\alpha)/2} + (\frac{1+\alpha}{1-\alpha})^{-(1+\alpha)/2}}.$$

Then f is strongly starlike of order α. Moreover, f is strongly convex of order $((1-\gamma)\alpha + \delta)/\gamma$.

Proof. We obtain f is strongly starlike of order α if we take $h(z) = zf'(z)/f(z)$ and $\alpha = \beta$ in Theorem 3.10. Since

$$\begin{aligned} &\left| \arg \left(1 + \frac{zf''(z)}{f'(z)} \right)^{\gamma} \right| - \left| \arg \left(\frac{zf'(z)}{f(z)} \right)^{1-\gamma} \right| \\ &\leq \left| \arg \left\{ \left(\frac{zf'(z)}{f(z)} \right)^{1-\gamma} \left(1 + \frac{zf''(z)}{f'(z)} \right)^{\gamma} \right\} \right| < \frac{\delta\pi}{2}, \end{aligned}$$

then by Theorem 3.10 we have

$$\begin{aligned} \left| \arg \left(1 + \frac{zf''(z)}{f'(z)} \right)^{\gamma} \right| &\leq \left| \arg \left(\frac{zf'(z)}{f(z)} \right)^{1-\gamma} \right| + \frac{\delta\pi}{2} \\ &< \frac{(1-\gamma)\alpha\pi}{2} + \frac{\delta\pi}{2}. \end{aligned}$$

Hence f is strongly convex of order $((1-\gamma)\alpha + \delta)/\gamma$. ∎

References

[1] A. Ebadian and J. Sokół, On the subordination and superordination of strongly starlike functions, Math. Slovaca **66**, no. 4, 815–822 (2016).

[2] S. S. Kumar and P. Yadav, On Oblique Domains of Janowski Function. arXiv preprint arXiv:2104.08533 (2021).

[3] S. S. Miller and P. T. Mocanu, *Differential subordinations*, Monographs and Textbooks in Pure and Applied Mathematics, **225**, Marcel Dekker, Inc., New York (2000).

[4] W. Janowski, Some extremal problems for certain families of analytic functions. I, Ann. Polon. Math. **28**, 297–326 (1973).

[5] M. Nunokawa, S. Owa, H. Saitoh, N. E. Cho and N. Takahashi, Some properties of analytic functions at extremal points for arguments, (preprint, 2002).

[6] M. Nunokawa and J. Sokół, Strongly gamma-starlike functions of order alpha, Ann. Univ. Mariae Curie-Skłodowska Sect. A **67**, no. 2, 43–51 (2013).

[7] M. Nunokawa and J. Sokół, Certain results for the class of Carathéodory functions, Mediterr. J. Math. **17**, no. 1, Paper No. 11, 9 pp (2020).

[8] M. Nunokawa and D. K. Thomas, On convex and starlike functions in a sector, J. Austral. Math. Soc. Ser. A **60**, no. 3, 363–368 (1996).

[9] V. Ravichandran, F. Rønning and T. N. Shanmugam, Radius of convexity and radius of starlikeness for some classes of analytic functions, Complex Variables Theory Appl. **33**, no. 1-4, 265–280 (1997).

[10] V. Ravichandran, Y. Polatoglu, M. Bolcal and A. Sen, Certain subclasses of starlike and convex functions of complex order. Hacet. J. Math. Stat. **34**, 9–15 (2005)

[11] J. Sokół and L. Trojnar-Spelina, On a sufficient condition for strongly starlikeness, J. Inequal. Appl. **383**, 11 pp (2013).

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India
Email address: poojayv100@gmail.com

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India
Email address: spkumar@dce.ac.in

2021 4th International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE) | 978-1-6654-1429-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/RDCAPE52977.2021.9633719

ANN and PSO based Approach for Solar Energy Forecasting: A Step Towards Sustainable Power Generation

Md Tabish Ansari
Department of Electrical Engineering
Delhi Technological University
Delhi, India
authentictabish@gmail.com

M. Rizwan
Department of Electrical Engineering
Delhi Technological University
Delhi, India
rizwan@dce.ac.in

***Abstract*—For proper designing and development of solar photovoltaic system, forecasting of solar energy becomes very important. Power developed by the solar photovoltaic system depends upon meteorological parameters like temperature solar irradiance. Variation in these parameters causes variation in power generated by the photovoltaic system. Hence forecasting becomes important. In this paper solar energy forecasting is done using artificial neural network and particle swarm optimization based artificial neural network. Artificial neural network is well established method used for forecasting purpose. However there output can further be improved by applying optimization technique. Here particle swarm optimization technique is used MATLAB software is used for coding the optimized neural network and 'nftool' application is used for simple artificial neural network. For ANN percentage error comes out to be 4.18% and for ANN-PSO it comes out to be 3.23%**

Keywords—Solar energy, Forecasting, Artificial Neural Network, Optimization, Particle Swarm Optimization.

I. INTRODUCTION

Excessive use of fossil fuel is causing pollution of environment lading to global warming which are creating challenges for the society. As a result use of renewable source of energy for generation of power are becoming more popular. Among various source of renewable energy such as wind energy, solar energy, geothermal energy and hydro energy. Solar energy is the most popular as it is readily available [1]. As a result solar energy based photovoltaic (PV) system are becoming popular. However power generated by the solar photovoltaic system (SPV) mainly depends upon solar irradiance and temperature. Fluctuation in these parameters causes fluctuation in the power generated by solar photovoltaic system. Variation in these parameters are unpredictable. Solar irradiance depends upon cloud cover, wind speeds, time of day etc. Therefore forecasting of solar PV power becomes important for proper designing of SPV system. For proper forecasting of solar power, forecasting of solar irradiance becomes very important [2].

Literature review related to solar energy forecasting is done. Forecasting of global solar energy and PV power using artificial neural network (ANN) by taking into consideration various meteorological parameters is discussed in [1]. Multi linear regression along with fuzzy logic incorporating various atmospheric condition for forecasting of solar irradiance is discussed in [2]. Solar radiation and solar power forecasting using fuzzy logic and ANN is discussed in [3], [4]. Short term global solar irradiance forecasting using seasonal auto regressive and nonlinear auto regressive based on ANN theory is discussed in [5]. Short term solar energy forecasting and the data used in this paper is discussed in [6]. SPV power forecasting is done by training neural network by Extreme Learning Machine (ELM) algorithm and is optimised by Particle Swarm Optimisation (PSO) technique is discussed in [7] and Genetic Algorithm and ANN is used for forecasting SPV power is discussed in [8]. Comparative study among ANN optimised by various advance algorithm like GA, PSO etc. is discussed in [9]. Training of neural network using PSO is discussed in [10]. Different types of advance algorithm which are used for optimisation purpose is discussed in chapter 2 of [11]. In this paper forecasting of solar energy is done using ANN and ANN optimised by PSO. Models developed are simple and optimised ANN give better results.

In this paper, basics of artificial neural network and ANN model developed is discussed in Section II. ANN results are presented and discussed in Section III. Basics about optimisation and Particle Swarm Optimisation is discussed is discussed in Section IV. In Section V ANN-PSO results are discussed. In Section VI comparison is done between ANN and ANN-PSO result. Then Section VII is conclusion and is followed by the reference.

II. ARTIFICIAL NEURAL NETWORK

A. Artificial neural network concept

Artificial neural network is computing system which operates similar to the biological neural network of human brain. Artificial neural network can be said as the interconnection of nonlinear neuron [3]. They can perform various task such as pattern recognition, optimization, forecasting, prediction etc. ANN can be used for both linear and nonlinear system. Relationship between input and output can easily be found by using ANN irrespective of the system. Generally ANN consist of three layers [4]. Input layer (data is received by this layer) is the first layer, hidden layer (this layer process the data) is the second layer and output layer (this layer receives the processed data from hidden layer) is the third layer. Fig.1 show the basic structure of ANN.

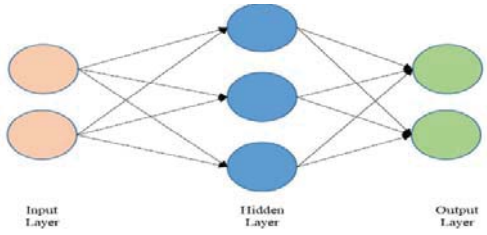


Fig. 1. Basic structure of ANN

As previously discussed that ANN has three layers. However number of hidden layer and number of neurons required for each layer is different for different system. Further weights and biases are applied to the inputs, weights

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:03:38 UTC from IEEE Xplore. Restrictions apply.

decides how much the input will affect the output and biases ensures that neurons are active even when input is zero. Major portion of data fed to an ANN model is used for training the network while smaller equal portion of data is used for testing and validation check. Fig.2 shows how weights and biases are associated with the input.

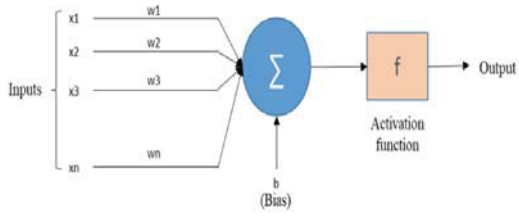


Fig. 2. Weights and biases in ANN

B. Methodology

Artificial neural network model is developed for forecasting of solar irradiance by using 'nftools' application provided in MATLAB software. Data for the forecasting purpose is taken from [6]. Data consist of solar irradiance measured at five different time period (time is taken in 24 hour clock system) and is shown in Table.1. Four inputs representing solar irradiance measured at four time periods i.e 09:00, 10:00, 11:00, 12:00 are provided. Using the ANN model, forecasting is done for the 13:00. Number of neurons taken here are 15. Training of artificial neural network is done by using Levenberg-Marquardt algorithm. This is the fastest back propagation algorithm available. In this paper 70% of data is used for testing the ANN model, while 15% of data is used for testing the model developed and 15% of the data is used for the validation check of the model developed. Fig.3 shows model developed by using nftool application of MATLAB.

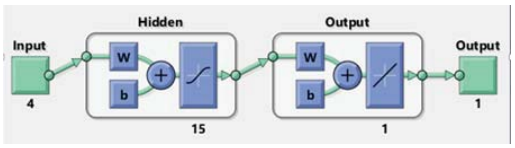


Fig. 3. ANN model developed using nftool

In this paper multilayer feed forward neural network (FFNN) is developed as it is most popular. FFNN does not have any feedback path [3]. The data flows in the forward direction only. Multilayer network topology is generally used for studying linear/nonlinear relationship between input and output. Flow chart showing the functioning of ANN is shown in Fig.4.

TABLE. 1. IRRADIANCE DATA FOR OCTOBER MONTH

Input				Output
09:00	***10:00***	***11:00***	***12:00***	***13:00***
0.848	0.786	0.812	0.858	0.898
0.858	0.808	0.882	0.676	0.646
0.716	0.785	0.845	0.863	0.796
0.817	0.808	0.877	0.875	0.580
0.784	0.741	0.720	0.830	0.823
0.817	0.708	0.810	0.753	0.803
0.731	0.740	0.793	0.376	0.732
0.696	0.638	0.803	0.800	0.741
0.857	0.852	0.887	0.882	0.841
0.877	0.854	0.896	0.871	0.830
0.837	0.780	0.800	0.781	0.776
0.836	0.789	0.772	0.780	0.733
0.819	0.745	0.762	0.365	0.381
0.100	0.100	0.100	0.100	0.100
0.739	0.730	0.766	0.737	0.653
0.850	0.752	0.793	0.815	0.750
0.809	0.702	0.779	0.772	0.712
0.734	0.747	0.809	0.805	0.772
0.795	0.764	0.781	0.767	0.685
0.703	0.650	0.718	0.693	0.559
0.674	0.717	0.707	0.710	0.584
0.602	0.630	0.695	0.713	0.603
0.618	0.665	0.729	0.650	0.583
0.567	0.599	0.658	0.638	0.549
0.486	0.490	0.596	0.577	0.480
0.570	0.581	0.672	0.669	0.572
0.553	0.578	0.698	0.698	0.626
0.567	0.580	0.689	0.667	0.561
0.430	0.665	0.134	0.417	0.498
0.472	0.506	0.479	0.683	0.550
0.412	0.520	0.569	0.481	0.458

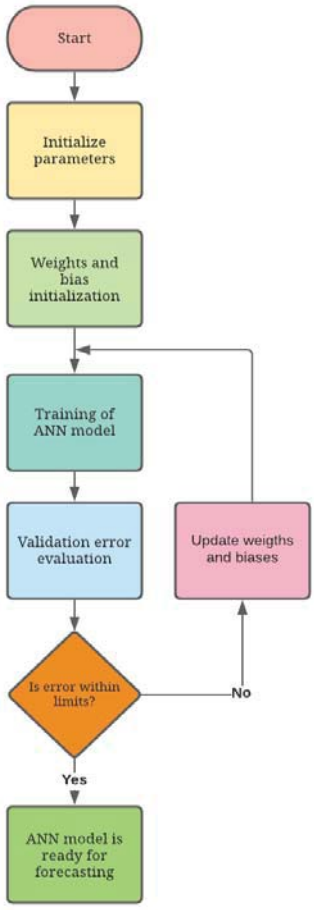


Fig. 4. Flow chart showing functioning of ANN

III. ANN RESULTS

Data for the development of model is taken from [6]. Data is for the month of October and is shown in Table 1. Data is in normalised form. Equation 1 is used for the converting actual data into normalized form. The result obtained is tabulated in Table 2. Equation 2 shows the formula for the mean absolute error. Fig.5 shows the validation performance plot for the ANN model. It shows how the ANN model is performing by measuring mean squared error. Fig.6 shows the regression graph obtained for the ANN model. It shows how much the forecasted data is close to actual data. Fig.7 shows ANN forecasted data and actual data compared. Solar irradiance in fig.7 is given in W/m^2

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:03:38 UTC from IEEE Xplore. Restrictions apply.

$$Y_s = \frac{(A_{max}-A_{min})}{(Y_{max}-Y_{min})}(Y - Y_{min}) + Y_{min} \quad (1)$$

Y_s = Normalized data.

Y = Original data.

Y_{max} = Maximum value in a data set.

Y_{min} = Minimun value in a data set.

A_{max} = Upper limit (0.9)

A_{min} = Lower limit (0.1)

$$\frac{1}{x}\sum_{t=1}^{n}\left|\frac{A_t-B_t}{A_t}\right| \quad (2)$$

A_t = Actual Value.

B_t = Forecasted Value

x = Number of days

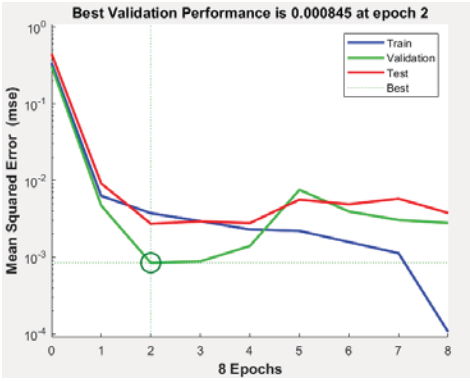


Fig. 5. Graph showing ANN model performance

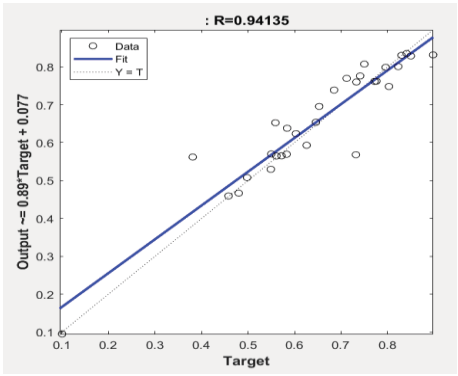


Fig. 6. Graph showing regression for the ANN model

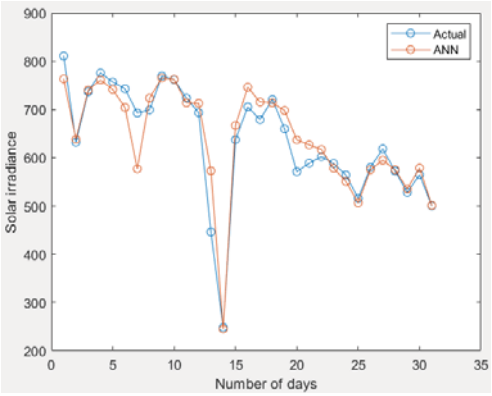


Fig. 7. Actual data and ANN data compared

IV. Optimization

A. Basic Concept

Optimization can simply be defined as the process by which best result for a particular problem can be obtained under given condition. In optimization the condition at which function gives maximum or minimum value can be found easily. Main purpose of optimization is to minimize the effort and maximize benefits. Earlier optimization involved long calculation and guess work to predict the data [7], [8]. But now with the development of the high speed digital computers, complex optimization process are possible.

Five steps are involved in development of optimization model. Firstly, the data is to be collected. Secondly, formulate and define the problem to be solved. Thirdly, developed the model as per the problem. Fourthly, validate the model and evaluate its performance and lastly, after performance evaluation interpret the result obtained.

Optimization of large scale problems involve several difficulties such as differentiability, dimensionality etc. [11]. Linear programming and dynamic programming are some conventional technique which fail to solve these problem. To overcome these difficulties, advance optimization techniques were developed based the evolution, behavior of animals. These algorithm are known as population based algorithm. Some of the famous population based algorithm are: Genetic Algorithm (GA) based on survival of the fittest and evolution theory. Particle Swarm Optimization (PSO) based on behavior of swarm of birds foraging. Ant Colony Optimization (ACO) based on behavior of ant foraging for food. Other optimization techniques are Harmony Search (HS), Artificial Bee Colony (ABC), and Grey Wolf Optimization (GWO) [11]. In this paper, PSO technique is used to improve the result obtained from the previous ANN model.

B. Particle Swarm Optimization

Particle swarm optimization algorithm is based on the behaviour of swarm of birds or fish. This optimization algorithm was developed by Eberhart and Kennedy [9]. It is evolutionary computation technique and has several advantage. In PSO algorithm 'particles' are considered as solution of problem. Particles are capable of moving in search landscape (Search landscape consist of all possible input and their corresponding objective values). Each particle has a best solution associated with it, and is known as 'pbest'. PSO also track global best values and is known as 'gbest' [10]. PSO involves continuously changing velocity of each particle so that the particles move towards its best location. In PSO algorithm velocity and position of particle is continuously updated till optimum value of objective function is reached.

$$V_{t+1} = wV_t + c_1r_1(P_t - X_t) + c_2r_2(G_t - X_t) \quad (1)$$

$$X_{t+1} = X_t + V_{t+1} \quad (2)$$

Equation 1 represents the velocity vector and equation 2 represent the position vector. Both these vectors are required for accurate determination of position of each particle. From equation 1 intensity of movement of particle can be determined. From equation 2 particle position in landscape can be determined. Here r_1 and r_2 are randomly generated number between 0 and 1, c_1 and c_2 are acceleration constant and w is inertia weight [11], [12], [13]. Fig.8 show the flow chart explaining the functioning of PSO algorithm

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:03:38 UTC from IEEE Xplore. Restrictions apply.

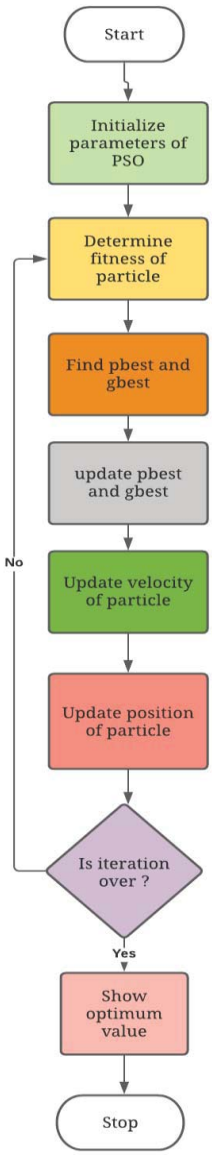


Fig. 8. PSO algorithm flow chart

V. PSO OPTIMIZED ANN RESULT

Whenever ANN is used, during training of ANN model weights and biases are randomly assigned. It changes whenever the ANN model is trained. When PSO algorithm is applied to ANN, the algorithm finds optimum value of weights and biases.

The objective function for PSO algorithm is chosen as mean square error (training error) [9]. The purpose of PSO in this paper is to minimise the mean square error (training error) by finding the optimum value of weights and biases. Initially the algorithm will start with random weights and biases and will determine the position and velocity for each iteration and determine the value of objective function and will keep on updating the pbest and when best value of objective function is obtained it is finally saved as gbest. Here pbest represent lowest error and gbest represent overall lowest error. Once optimum value of weights and biases are obtained ANN model is trained again with the updated weights and biases and the results are obtained and shown in Table 2. Number neurons are 15. Number of particles taken are 10 and maximum number of iteration is 1000. Data for the ANN-PSO model is shown in Table.1.

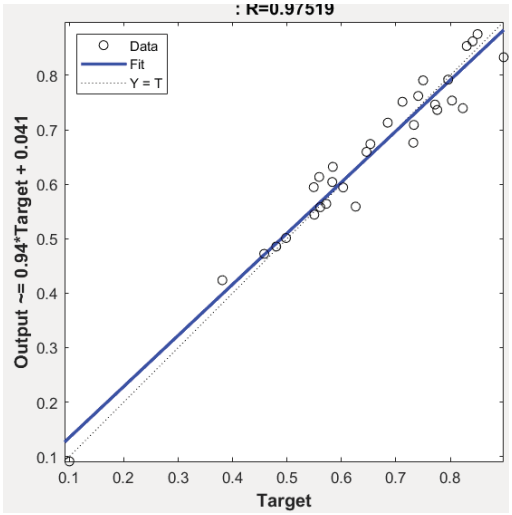


Fig. 9. ANN-PSO regression graph

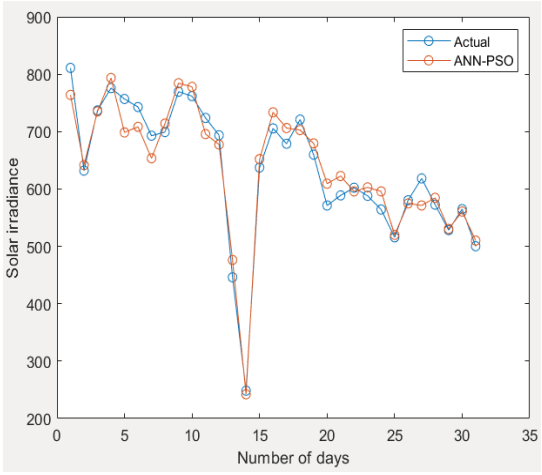


Fig. 10. Graph comparing actual and ANN-PSO data

Fig.9 shows the regression plot PSO optimized ANN. From this graph we can see that the forecasted data is much closer to the actual data as compared to ANN. Fig.10 show the comparison of the actual data and the data obtained from ANN-PSO model (solar irradiance is given in W/m^2).

VI. COMPARISON OF ANN AND ANN-PSO RESULT

For ANN model, from the Table.2 it is observed that minimum absolute relative percentage error (ARE) comes out to be 0.1% and maximum absolute relative error comes out to be 28.52%. Mean absolute percentage error comes out to be 4.18% for fuzzy logic model.

TABLE. 2. Output of ANN and ANN-PSO

Actual Data (W/m^2)	ANN Output (W/m^2)	ARE in ANN output (%)	ANN-PSO output (W/m^2)	ARE in ANN-PSO output (%)
811	763.2	5.8	764	5.8
632	638.03	0.95	641.5	1.5
737	740	0.4	735.1	0.25
775.84	761.1	1.9	793.4	2.3
756.86	741.4	2.04	698.5	7.72
742.79	704.1	5.21	708.3	4.64
692.87	577	16.72	653.5	5.68

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:03:38 UTC from IEEE Xplore. Restrictions apply.

699.2	723.81	3.46	714	2.1
769.51	765.3	0.58	784.3	1.8
761.78	762.5	0.1	778	2.1
723.81	714	1.34	695.7	3.8
693.58	713	2.88	677.4	2.33
446.07	573	28.52	476.3	6.7
248.5	245	1.4	242.2	2.5
637.33	667	4.56	652.1	2.3
705.53	746.1	5.78	733.6	3.9
678.81	715.6	6	706.2	4
721	714	0.91	702.7	2.5
659.83	698	5.78	679.5	2.9
571.23	637	11.56	609.2	6.2
588.81	626.7	6.44	622.6	5.74
602.17	617	2.5	595.8	1.05
588.11	578.3	1.7	602.8	2.5
564.20	550.8	2.37	595.8	5.6
515.68	506.5	1.78	520	0.83
580.37	575.4	0.86	574.8	0.96
618.34	595.1	3.75	571.2	7.6
572.64	575.4	0.48	584.6	2.1
528.34	535.4	1.14	530.4	0.4
564.91	578.9	2.35	560.7	0.7
500.22	501	0.18	510.1	1.26

For ANN-PSO model, from Table.2 it is observed that the minimum absolute relative error comes out to be 0.25% and maximum absolute relative error comes out to be 7.72% and mean absolute error comes out to be 3.23%. Fig.11 shows both the ANN result and Fuzzy result and are compared to the actual. Here solar irradiance in fig.11 is given in W/m^2.

VII. Conclusion

Forecasting of solar energy is done using ANN and ANN optimized by PSO algorithm and mean absolute percentage error (MAPE) is calculated for both the models. For ANN model it MAPE comes out to be 4.18% and for ANN-PSO model MAPE comes out to be 3.23%. Percentage error for both the models is quite low and both the models give accurate result. However the PSO optimized ANN model has lower error as compared to simple ANN model. Thus by applying the optimizing technique on the ANN model the performance of the ANN model improves. Thus accuracy is improved and better result are obtained. Once the solar energy forecasting is done, accurate forecasting of SPV system power output can be done and hence effective and efficient designing of SPV system is possible. These forecasting technique can also be used for the efficient development of maximum power point tracker. Further reliability analysis of SPV system can be also be done.

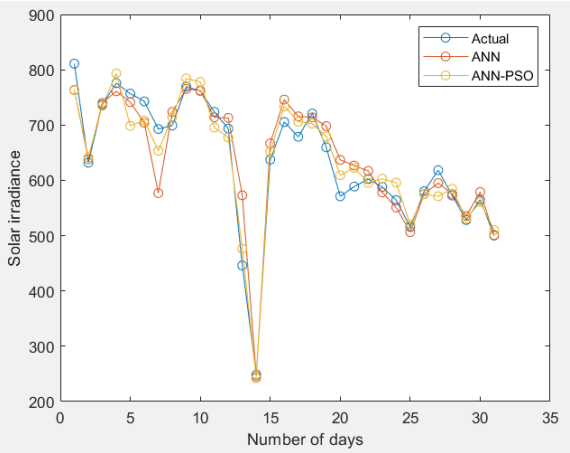


Fig. 11. Graph comparing actual data with ANN and ANN-PSO

References

[1] Gulnar Perveen, M. Rizwan, Nidhi Goel & Priyanka Anand (2020) Artificial neural network models for global solar energy and photovoltaic power forecasting over India, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, DOI: 10.1080/15567036.2020.1826017

[2] Mehta, Sahil & Basak, Prasenjit. (2019). Solar irradiance forecasting using fuzzy logic and multilinear regression approach: a case study of Punjab, India. International Journal of Advances in Applied Sciences. 8. 125-135. 10.11591/ijaas.v8i2.pp125-135

[3] S.X. Chen, H.B. Gooi, M.Q. Wang, "Solar radiation forecast based on fuzzy logic and neural networks", Renewable Energy, Volume 60, 2013, Pages 195-201, ISSN 0960-1481, https://doi.org/10.1016/j.renene.2013.05.011

[4] Z. P. Ncane and A. K. Saha, "Forecasting Solar Power Generation Using Fuzzy Logic and Artificial Neural Network," 2019 Southern African Universities Power Engineering Conference/Robotics and Mechatronics/Pattern Recognition Association of South Africa (SAUPEC/RobMech/PRASA), Bloemfontein, South Africa, 2019, pp. 518-523, doi: 10.1109/RoboMech.2019.8704737

[5] A. D. Orjuela-Cañón, J. Hernández and C. R. Rivero, "Very short term forecasting in global solar irradiance using linear and nonlinear models," 2017 IEEE Workshop on Power Electronics and Power Quality Applications (PEPQA), Bogota, 2017, pp. 1-5, doi: 10.1109/PEPQA.2017.7981691.

[6] A. Chugh, P. Chaudhary and M. Rizwan, "Fuzzy logic approach for short term solar energy forecasting," 2015 Annual IEEE India Conference (INDICON), New Delhi, India, 2015, pp. 1-6, doi: 10.1109/INDICON.2015.7443206.

[7] Manoja Kumar Behera, Irani Majumder, Niranjan Nayak, Solar photovoltaic power forecasting using optimized modified extreme learning machine technique, Engineering Science and Technology, an International Journal, Volume 21, Issue 3, 2018, Pages 428-438, ISSN 2215-0986, https://doi.org/10.1016/j.jestch.2018.04.013.

[8] Pattanaik, Debasish, Mishra, Sanhita, Khuntia, Ganesh Prasad, Dash, Ritesh and Swain, Sarat Chandra. "An innovative learning approach for solar power forecasting using genetic algorithm and artificial neural network" *Open Engineering*, vol. 10, no. 1, 2020, pp. 630-641. https://doi.org/10.1515/eng-2020-0073

[9] Le, L.T.; Nguyen, H.; Dou, J.; Zhou, J. A Comparative Study of PSO-ANN, GA-ANN, ICA-ANN, and ABC-ANN in Estimating the Heating Load of Buildings' Energy Efficiency for Smart City Planning. *Appl. Sci.* **2019**, *9*, 2630. https://doi.org/10.3390/app9132630

[10] H. T. Rauf, W. H. Bangyal, J. Ahmad and S. A. Bangyal, "Training of Artificial Neural Network Using PSO With Novel Initialization Technique," 2018 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), 2018, pp. 1-8, doi: 10.1109/3ICT.2018.8855743.

[11] Rao R.V., Savsani V.J. (2012) Advanced Optimization Techniques. In: Mechanical Design Optimization Using Advanced Optimization Techniques. Springer Series in Advanced Manufacturing. Springer, London. https://doi.org/10.1007/978-1-4471-2748-2_2

[12] Aljanad, A.; Tan, N.M.L.; Agelidis, V.G.; Shareef, H. Neural Network Approach for Global Solar Irradiance Prediction at Extremely Short-Time-Intervals Using Particle Swarm Optimization Algorithm. Energies 2021, 14, 1213. https://doi.org/10.3390/en14041213

[13] Ibnelouad, A, El Kari, A, Ayad, H, Mjahed, M. Improved cooperative artificial neural network-particle swarm optimization approach for solar photovoltaic systems using maximum power point tracking. Int Trans Electr Energ Syst. 2020; 30:e12439. https://doi.org/10.1002/2050-7038.12439

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:03:38 UTC from IEEE Xplore. Restrictions apply.

The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/2053-4604.htm

Barriers to corporate social responsibility implementation in the medium size manufacturing sector: an interpretive structure modelling approach

Gaurav Dawar
Department of Humanities, Delhi Technological University, Delhi, India and Christ University – Delhi NCR Campus, Ghaziabad, Uttar Pradesh, India, and
Seema Singh
Department of Humanities, Delhi Technological University, Delhi, India

Received 20 December 2020
Revised 18 April 2021
5 July 2021
8 September 2021
10 November 2021
Accepted 11 November 2021

Abstract

Purpose – Corporate social responsibility (CSR) practices are gaining momentum globally but their implementation becomes problematic due to the presence of barriers. So, this study aims to identify the barriers to CSR implementation among manufacturing enterprises, develop their classification and establish relationships among the barriers.

Design/methodology/approach – An exhaustive list of barriers was identified from the literature, and following surveys and expert opinions, 19 critical barriers were extracted. Interpretive structure modelling was used to understand the hierarchal and contextual relationships among barriers of CSR implementation.

Findings – The results show that are no autonomous variables present in the study. The proposed conceptual framework presents the hierarchy and interlinkage of barriers to CSR implementation in manufacturing enterprises. The results also indicate that rigidity in culture and corruption in the system and within the governance system of the country are the two most influential barriers that impede CSR implementation in manufacturing enterprises.

Originality/value – The interactions among CSR barriers provide policymakers, industrial practitioners and managers with a framework to recognise and evaluate mutual relationships and interlinking among barriers. CSR training and undertaking CSR in collaboration can help medium enterprises overcome these barriers and prepare strategies to mitigate their impact.

Keywords Manufacturing industries, Barriers, Interpretive structural modelling, Corporate social responsibility, Medium-sized enterprises, CSR Barriers, Manufacturing enterprises

Paper type Research paper

1. Introduction

Barriers to corporate social responsibility (CSR) implementation are obstacles preventing implementing a policy instrument that leads to a less effective social strategy. These barriers limit the strategic importance of CSR due to hurdles during implementation at the organisational level. These Barriers vary with enterprise size and their economic impact (Laudal, 2011; Álvarez Jaramillo *et al.*, 2019). The generic CSR barriers are moral imperatives of enterprises, the relationship of the owner's personality trait and responsible behaviour, resource access and external environment (Lepoutre and Heene, 2006). An effective solution

Journal of Entrepreneurship in Emerging Economies
2053-4604
DOI 10.1108/JEEE-12-2020-0451
© Emerald Publishing Limited

is required for enterprises to balance their social performance towards sustainable development while enhancing their competitiveness.

CSR is a concept that denotes corporates' voluntary contribution to a better society including a cleaner environment (European Commission, 2001). Worldwide growing concern towards CSR in the previous two decades has motivated companies to demonstrate their social concern and create shared value (Stoian and Gilman, 2017). Enterprises CSR orientation can help them attain holistic and all-inclusive organisational culture by inculcating progressive, high performance and innovative organisation [1]. The demonstration of social concern can generate capabilities to survive in a competitive, complex and dynamic environment by focussing on stakeholders' expectations (Solano *et al.*, 2016). These expectations are no longer limited to large companies or multinational corporations but also to small- and medium-sized enterprises (SMEs), as entities using societal resources can never act as an isolated economic entity, they should understand, develop and work for the welfare of society (Porter and Kramer, 2006). Research on SMEs engagement in CSR has recently started to emerge (Morsing and Perrini, 2009; Fassin, 2008). The previous works on CSR implementation have been undertaken in developed economies including Denmark, Hungary, Greece, UK, Germany and the USA (Metaxas and Tsavdaridou, 2010; Castka *et al.*, 2004; Chen and Bouvain, 2009) and developing economies such as India and Cameroon (Gautam and Singh, 2010; Demuijnck and Ngnodjom, 2013 Ward *et al.*,2007). There is a lack of theoretical and empirical investigations for CSR implementation in developing countries (Graafland and Smid, 2019; Arevalo and Aravind, 2011) and the current study addresses this gap. The evolving nature of the CSR concept and complex business environment in countries in developing economies create distinct barriers that hamper CSR's effective implementation (Bux *et al.*, 2020). The literature from developed and developing economies has addressed the barriers to CSR implementation from a sectoral perspective, such as; the textile industry (Shen *et al.*, 2015); supply chains (Faisal, 2010), shipping (Yuen and Lim, 2016), construction (Zhang *et al.*, 2019), airlines (Kuo *et al.*, 2016), food sectors (Hartmann, 2011), pharmaceutical industries (Min *et al.*, 2017), hotel industry (Garay and Font, 2012), automobile sector (Luthra *et al.*, 2011), service industry (Kumar *et al.*, 2019a, 2019b) and the manufacturing industry (Goyal and Kumar, 2017; Bux *et al.*, 2020). The literature has highlighted the barriers faced by large enterprises are lack of training, limited financial resources and difficulty obtaining information about CSR implementation (Arevalo and Aravind, 2011), whereas smaller enterprises experience constraints to financial resources, human resources and time as well as low cost to benefit ratio as barriers to CSR implementation (Sweeney, 2007; Laudal, 2011). However, the barriers to medium-sized manufacturing enterprises are still underexplored. Rentizelas *et al.* (2020) emphasised regular studies on barriers to CSR implementation so that enterprises continually prioritise them and improve enterprises' social performance to cope with the evolving nature of the concept.

Medium size manufacturing enterprises distinguish themselves from large enterprises for turnover, investment in assets, market share, number of employees and ownership structures (Curran and Blackburn, 2000). They are regarded as a place for the continuous improvement of products, processes and technologies (Chikara and Sahay, 2008). Generally, small and medium enterprises are grouped together for CSR research with limited distinction or focus on medium-sized enterprises' CSR practices. The previous works (Lamberti and Noci, 2012; Preuss and Perschke, 2010; Murillo and Lozano, 2006; Lyons, Bartlett and McDonald, 2016) highlight that CSR activities of medium-sized manufacturing enterprises pose unique challenges as the concept is nuance, complex and executed at a scale not large enough to address its own challenges (He *et al.*, 2014; Álvarez Jaramillo *et al.*, 2019).

The CSR activities in medium enterprises are highly influenced by the value system of owners/managers and restricted to local surroundings (Refer to Appendix for the distinction between small, medium and large enterprises). Other factors that influence the CSR implementation of manufacturing enterprises are global institutional pressure and enterprises' operating environment, experiential learning of managers and stakeholder expectations (Lyons *et al.*, 2016; Waddock and Graves, 1997).

The current study on CSR implementation is important as CSR implementation in many Asian countries is in an early stage and requires further exploration (Hsu and Cheng, 2012; Lee *et al.*, 2012). The significant global contribution to gross domestic product and employment by medium enterprises (Raynard and Forstater, 2002) has led the governments to acknowledge that the CSR movement will be incomplete without medium enterprises' active engagement in CSR (Kechiche and Soparnot, 2012; Spence, 1999). Despite, these facts, a handful of studies address CSR barriers in medium-size enterprises in emerging economies (Goyal and Kumar, 2017; Arevalo and Aravind, 2011; Kumar *et al.*, 2019a, 2019b) and the current study suggests alternatives to mitigate those challenges. Another important motive for this study is that focussing on CSR practices of medium enterprises in an emerging economy can provide scholars and practitioners with a new dimension of social responsibility implementation and inform about the attributes not discussed earlier in the sustainable development literature. Thus, the study examines the barriers to CSR implementation in medium-sized manufacturing enterprises in an emerging economy.

Next section discusses the literature on barriers to CSR in manufacturing enterprises, followed by a discussion of a survey of owners/managers of 121 medium-sized enterprises including enterprises of international origin to identify the portfolio of barriers. The responses are analysed using interpretive structure modelling (ISM) and the significance of barriers is based on Matrice d'Impacts croises-multiplication appliqué a classment (MICMAC) analysis to classify them based on its power. This is followed by implications and a conclusion.

2. Review of literature

2.1 Theoretical background of barriers to corporate social responsibility implementation

Developing corporate policy is just the first step towards goal achievement and achieving desired results requires policies' effective implementation. The challenges and barriers faced by the enterprises in the implementation of policy are referred to as "Implementation barriers" (Bach-Mortensen *et al.*, 2018; Orlitzky *et al.*, 2011). The barriers are interdependent and understanding how the barriers are interconnected can help managers design an optimal solution to the problems they face in CSR implementation (Goyal and Kumar, 2017). The barriers to CSR, in general, have been explored (Vives, 2005; Dearing, 2000). For example, Carroll's CSR Pyramid (1991) identified many factors that influence CSR implementation in organisations. Despite his work's relevance, it was criticised for prioritising economic responsibility (Key, 1999; Heugens and Oosterhout, 2008) which corresponds to Friedman's (19070) controversial editorial "Social Responsibility of business is to increase its profit". In line with the work of Carroll (1991), Arevalo and Aravind (2011) examined the CSR performance of Indian companies participating in the UN Global Compact (UNGC). The study catered to a limited number of companies who were participants at UNGC and the practices of non-participants are ignored. One of the most prominent factors that affect CSR implementation is economic conditions (Campbell, 2007) and maximisation of shareholder value (Arevalo and Aravind, 2011; Jones, 1996). In addition, limited

availability of resources such as financial, human and time-related and firm size are key influencers of CSR implementation (López-Gamero *et al.*, 2008; Muller and Kolk, 2010; Gamerschlag *et al.*, 2011).

The literature has highlighted different perspectives for categorising CSR implementation barriers based on managers' understanding of CSR. Bello and Kamanga (2020) categorised barriers as internal and external barriers from the perspective of the organisation's stakeholder pressure. Lack of resources, lack of awareness and knowledge about CSR are internal barriers and government support, ambiguity in CSR policy, failure to cooperate with peers and mismanagement of CSR resources by communities are categorised as external barriers. Agudo-Valiente *et al.* (2017) categorised barriers based on the classic and contemporary approach of CSR implementation as subjective barriers and objective barriers. Subjective CSR barriers are complex to define and associated with managers and owners' values and beliefs, such as lack of ethical integration, CSR as an image strategy and lack of commitment to transparency and objective CSR barriers are outcome-oriented and based on organisational accomplishments such as lack of resources, low institutional interest and challenging to interpret CSR (Agudo-Valiente *et al.*, 2017).

Another classification of CSR in SMEs is moral-driven and profit-driven (Zientara, 2017; Govindasamy and Suresh, 2017), whereas in large enterprises CSR is categorised as explicit and implicit (Matten and Moon, 2008). Moral-driven influencers are related to morality, charity and honesty conditioned by moral beliefs and barriers whose recognition is not prepared by such faith (Agudo-Valiente *et al.*, 2017). Profit-driven influencers are related to undertaking CSR activities that would make a net contribution to profit due to managers' belief in earning money for owners and compliance with legislation (Greenfield, 2004). Managers holding such a view on CSR would tend to have less commitment towards CSR. Identifying these barriers may help managers formulate optimal strategies to mitigate them (Goyal and Kumar, 2017). Among these barriers, 19 barriers applicable to medium-sized enterprises were identified in this study. A brief description of these barriers is provided in Table 1.

2.2 Characteristics and eco-system of medium-size manufacturing enterprises

Manufacturing enterprises make a significant economic contribution (Chikara and Sahay, 2008) and leave significant environmental footprints (Preuss and Perschke, 2010). The definition of small, medium and large enterprises vary across economies, for example, Japan uses employment as a differentiator, Singapore uses Fixed Asset and employment as an indicator and India uses investment in fixed assets and turnover as the differentiating criteria (Dixit and Pandey, 2011; Hudson *et al.*, 2001). SMEs are not smaller versions of large-size enterprises and vice versa. Recent studies have found that these enterprises struggle to make a profit due to increased technology-based competition (Govidarajan *et al.*, 2019). The enterprises are flexible (Dex and Scheibl, 2001), informally managed (Coyte *et al.*, 2012) and dependent upon individual decision makers (Sherer, 2003). The multi-tasking owners and managers (Choi *et al.*, 2018) make these enterprises capable to adopt a fast-changing global market environment (Prasanna *et al.*, 2019). They feel the pressure to reform and improve stakeholder-related practices responsibly and profitably (Waddock and Graves, 1997). A source of pressure is adherence to Global Standards such as Global Reporting Initiatives (GRI), ISO-26000, which define the expected level of corporate responsibility and report economic, social and environmental performance (Waddock *et al.*, 2002). Hence, these organisations have no choice but to respond responsibly. Many medium-sized enterprises are reluctant to initiate CSR due to their inability to overcome these barriers

Medium size manufacturing sector

Table 1. Barriers to CSR

Perspective	Element no.	Barriers to CSR	Description of the barriers	References
Medium-sized enterprise perspective	C1	Lack of knowledge and expertise	There is an apparent lack of awareness and general knowledge among crucial decision makers about energy sources, causes of pollution and ways to address them. Lack of eco-literacy among SME managers and customers and ignorance of their environmental impact are barriers for SMEs to undertake environmental issues	Madsen and Ulhøi (1996) (DE); Shirodkar *et al.* (2018) (DEE); Puppim de Oliveira and Jabbour (2017) (DEE); Mzembe *et al.* (2016) (DEE); Coyle (2005) (DE); Seidel *et al.* (2009) (DE) and Yuen and Lim (2016) (DE)
	C3	Limited resources for delivering CSR programme	Management time and resources are limited and scarce resources may force enterprises to maintain their core activities and put social responsibility discharge at a halt	Yuen and Lim (2016) (DE); Kusyk and Lozano (2007) (DE) and Apaydin *et al.* (2020) (DE)
	C4	Unable to decide the best social use of CSR fund	The SMEs cannot model the CSR process of stronger companies due to resource constraints, shortage of skilled man force, lack of knowledge or non-availability of external partners may lead to deciding the best use of CSR funds	Pai (2018) (DEE); Mascarenhas (2018) (DEE); Sundar (2017) (DEE) and Rahman *et al.* (2020) (DEE)
	C5	Access to finance	SMEs have limited access to capital because of the higher cost of intermediation, lack of realistic financial valuation procedures and if they believe the cost, resources and time needed to implement CSR are high, it may lead to lower willingness to implement CSR	Cheng *et al.* (2014) (DE); Fitjar (2011) (DE); Jenkins (2006) (DE); Turyakira (2017) (DEE) and Agyemang and Ansong (2017) (DEE)
	C6	Company's internal value system or culture	Unfavourable attitude, work philosophy and behaviour of companies towards social and environmental change influence the value system of organisation and it may perceive green practices as an additional burden	Laudal (2011) (DE); Yuen and Lim (2016) (DE); Shen *et al.* (2015) (DEE); Chang *et al.* (2015) (DE) and Oreg (2003) (DE)
	C7	Lack of top management commitment	The unfavourable attitude of owners due to different risk perception and risk aversion attitude and resistance to change leads to conventional approach towards the business model and that may constituent as a major bottleneck in SMEs	Wilms *et al.* (1994) (DE); Chin *et al.* (2013) (DE); Arevalo and Aravind (2011) (DEE); Shen *et al.* (2015) (DEE) and Besch (2005) (DE)

(*continued*)

Table 1.

Perspective	Element no.	Barriers to CSR	Description of the barriers	References
	C17	Absence of organisational reward for employees for doing CSR	Rewards for Social performance promote personal satisfaction and employee welfare and the absence of rewards would reduce their motivation to involve in CSR projects	Montabon *et al.* (2016) (DE); Law *et al.* (2017) (DE); Cantor *et al.* (2012) (DE) and Bhattacharya *et al.* (2008) (DE)
Stakeholder perspective	C2	The low willingness of internal Stakeholders in participating and contributing to CSR activities	Due to lack of motivation arising due to the absence of incentive, the internal stakeholder is less willing to participate in CSR activities	Kusyk and Lozano (2007) (DE); Parviainen *et al.* (2018) (DE) and Lee *et al.* (2012) (DE)
	C10	Poor supplier and distributor commitment	Due to a lack of capability or capacity for sustainability management, supplier and distributor may have a poor commitment	Kumar *et al.* (2014) (DEE) and Villena and Gioia (2020) (DE)
	C12	Low involvement of external stakeholders	The external stakeholders are losing interest and low willingness to exchange information regarding involvement in responsible practices	Kusyk and Lozano (2007) (DE); Mathiyazhagan *et al.* (2013) (DEE); Baden *et al.* (2009) (DE) and Ciliberti *et al.* (2008) (DE)
	C13	Societal cynicism for CSR as a promotional activity	A society with a low attitude towards sustainability may perceive CSR cynically	Kuokkanen and Sun (2016) (DE); Jahdi and Acikdilli (2009) (DE); Ghazilla *et al.* (2015) (DEE) and Illia *et al.* (2013) (DE)
	C15	Lack of customer awareness	A less well-established CSR and lack of awareness keep customers uninformed and refrain from getting involved	Shen *et al.* (2015) (DEE) nd Luo *et al.* (2019) (DE)
Manufacturing industry perspective	C8	Lack of effective strategic planning for CSR	Lack of effective strategic planning for CSR can create CSR implementation obstacles that hamper its competitive position in the long run	Alotaibi *et al.* (2019) (DE); Faisal (2010) (DE) and Galbreath (2010) (DE)
	C18	Lack of external partner for CSR	Smaller organisations are dependent on partners' expertise in delivering CSR projects and their partners working under different environmental settings may find it difficult to match their CSR philosophy	Poret (2014) (DE); Husted (2003) (DE) and Faisal (2010) (DE)

(*continued*)

Table 1.

Perspective	Element no.	Barriers to CSR	Description of the barriers	References
	C11	Customers' low willingness to pay for CSR	The customer is price sensitive and concentrates more on economic profit rather than societal parameters. It motivates the manufacturer to focus on profit only	Yuen and Lim (2016), Shen *et al.* (2015) (DEE)
CSR attribute	C9	Absence of significant benefits for CSR implementation to enterprise	The absence of established literature on the relationship between CSR and firm performance creates a dilemma in the decision makers' minds	Cochran and Wood (1984) (DE); Janssen and Vanhamme (2015) (DE); Lee *et al.* (2012) (DE); Yoon and Chung (2018) (DE) and Wang *et al.* (2018) (DE)
	C19	Lack of benchmarking standards for CSR performances	Development of CSR learnings with no opportunity to benchmark or limited comparability would make little or no contribution to Sustainable development	Yuen and Lim (2016) (DE) and Shen *et al.* (2015) (DEE)
	C16	CSR implementation is too complex	The concept of CSR is multidisciplinary, due to which there is no single universally accepted definition. With the great variety of definitions and the fact that they have changed over time, CSR has been described as a "Chameleon concept". The amorphous nature of CSR slows down effective implementation	Jamali and Abdallah (2015) (DEE); Gupta *et al.*, 2017) (DE); Bhardwaj (2016) (DEE) and Maon *et al.* (2017) (DE)
Government policy	C14	Corruption	Corruption inside the company and in the regulatory environment creates a challenge for enterprises to obtain environmental certification and other clearance that hinders the implementation of social responsibility	Hasan (2016) (DEE); Hossain *et al.* (2016) (DEE) and He *et al.* (2014) (DE)

(Yuen and Lim, 2016; Bennington and Minutolo, 2013). Overcoming these barriers would provide these enterprises with distinctive capabilities (Aragón-Correa *et al.*, 2008) to remain competitive (Porter and Kramer, 2006) and improve their image (Jenkins, 2006). In comparison to large enterprises, medium enterprises are in a better position to engender commitment to their vision that makes them more socially responsible (Smith, 2013). These characteristics make medium enterprises unique and significant in the economy. Thus, there is a need to address CSR implementation barriers in medium-sized enterprises to understand their CSR practices.

2.3 Multi-criteria decision-making techniques in corporate social responsibility research

Multi-criteria decision-making (MCDM) is a term that was popularised by Stanley Zoints and is concerned with structuring and solving complex issues with multiple criteria for decision-making (Zionts, 1979). The increasing complexity of CSR issues and the limitation of one-factor analysis has necessitated the need for a multi-factor framework (Jamali and Karam, 2018). These techniques are gaining popularity in CSR research that enables the decision maker to consider all of the parameters related to problems that can be compromised and prioritise them (Kumar *et al.*, 2017). The MCDM techniques applied in CSR research are analytical hierarchy process (AHP), ISM, decision-making trial and evaluation laboratory technique (DEMANTEL) and data envelopment analysis (DEA) (Table 2).

2.4 Research gap

A proposition is that corporations in different sectors and countries face similar barriers while implementing CSR. However, the relative degree of importance associated with each barrier differs owing to the country-specific environmental factors. In recent years, researchers have attempted to identify the barriers related to CSR implementation in the manufacturing, supply chain and supplier selection process (Goyal and Kumar, 2017; Bux *et al.*, 2020; Faisal, 2010; Kumar *et al.*, 2014). However, further research is required to highlight the CSR implementation barrier in the manufacturing sector and the current study addressed this research gap. Previous studies have explored barriers to CSR implementation in large manufacturing enterprises. However, barriers to CSR implementation in medium enterprises are yet to be explored.

The ISM technique is applied to understand the contextual relationships among different variables. Further, these relationships can be tested through quantitative models to establish the relationships' validity and MICMAC analysis is used to classify barriers based on their strength and dependence.

3. Research methodology

In this study, quantitative and qualitative analyses are applied to address the question: "What are the barriers to CSR implementation in medium manufacturing enterprises"? The methodology adopted for this study is shown in Figure 1. For this, a questionnaire with 19 items on a 5-point Likert scale (1 is not a barrier to CSR implementation and 5 being a barrier of CSR implementation) was finalised in consultation with one academic and two industry experts who have been working in the area of social responsibility in manufacturing companies for more than two decades. The owners/senior-level managers involved in the social responsibility practices in medium enterprises were the targeted respondents. In total, 121 medium-sized manufacturing enterprises across five districts [2] were approached for the survey. The responses were collected in January-February 2020 with 87 usable

Sr.no	Author	Purpose	MCDM technique applied
1	Faisal (2010)	"Identified barriers to CSR in Supply Chain"	ISM-MICMAC
2	Shen *et al.* (2015)	"Identifying and analysing the barriers to CSR in the Indian Textile Industry"	AHP
3	Malek and Desai (2019)	"Prioritised the barriers to sustainable manufacturing in manufacturing organisations in India"	Best Worst Method
4	Kumar *et al.* (2019a, 2019b)	"Prioritised barriers to CSR in the Indian Service Industry"	AHP
5	Moktadir *et al.* (2018)	"Prioritised Drivers of CSR in the Footwear industry in Bangladesh"	Delphi and AHP
6	Bux *et al.* (2020)	"Identification of Barriers to CSR adoption in the manufacturing industry in Pakistan"	ISM
7	Raut *et al.* (2018)	"Identified and model critical barriers to sustainable practices in Oil and Gas Sector"	ISM
8	Goyal and Kumar (2017)	"Identify and rank barrier to CSR implementation in manufacturing companies"	ISM
9	Kumar *et al.* (2014)	"Analysing the CSR issues behind suppliers selection process in the Fireworks Industry in India"	ISM
10	Chang *et al.* (2015)	"Identification of factors for implantation of CSR in the Airline Industry"	DEMANTEL
11	Belu and Manescu (2013)	"Developing Strategic CSR index based on a multi-factor approach"	DEA
12	Kamble *et al.* (2018)	"Analysing the implementation barriers of dual cycling in port container terminal using interpretive structural modeling – Indian context"	ISM
13	Shibin *et al.* (2017)	"Drivers of Sustainable Supply Chain Management"	ISM

Table 2.
CSR research based on MCDM techniques

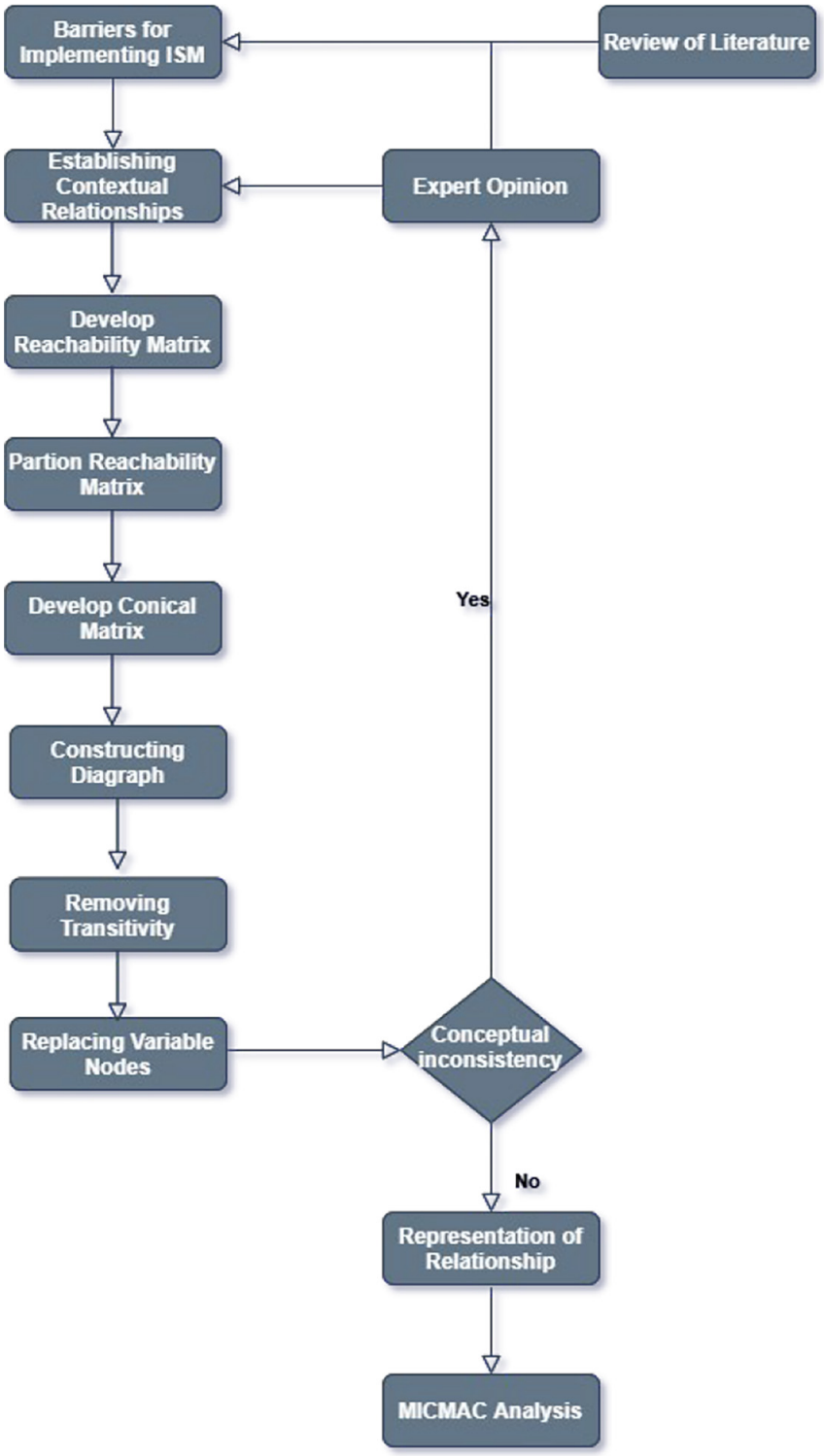


Figure 1.
Process of interpretive structure modelling

responses (i.e. a response rate of 72%) and overall Cronbach alpha of 0.86. Out of 87, 24% of responses were received from enterprises of international origin and 76% were from domestic enterprises. The response for barriers was analysed through the relative importance index (RII) using the following formula:

$$RII = \frac{\Sigma WiXi}{\Sigma Xi}, (0 \leq RII \leq 1)$$

where Wi is the constant presenting weight given to i; Xi is the variable presenting frequency of response for i = 1,2,3,4,5. The barriers with the RII above 0.4 and the mean above 3 are highly significant (Singh and Rathi, 2021). This resulted in only 15 of the 19 barriers being used for further investigation (Table 3). The ISM was then applied to prioritise barriers based on the degree of their impact on enterprises' CSR practices. The method is discussed in more detail in the next section.

4. Application of interpretive structure modelling

To identify the barriers faced by medium-sized manufacturing companies in implementing CSR, a total of 19 barriers applicable to social responsibility practices of manufacturing enterprises were identified from the literature (Table 1).

4.1 Interpretive structural modelling

Warfield's ISM (Warfiled, 1973) is a well-established technique for the portrayal of direct and indirect relationships among variables in multi-level structural modelling (Opricovic and Tzeng, 2003; Thakkar *et al.*, 2008). The ISM process starts with the identification of variables, further direct and indirect contextual relationships among the variables are identified, which are converted into structural self-interaction matrix (SSIM) and reachability matrix and finally a structural model through hierarchical iterations (Sushil, 2012).

Element no.	CSR barriers	Mean	RII
C1	Lack of knowledge and expertise	4.27	0.790
C2	The low willingness of internal stakeholders in participating and contributing to CSR activities	3.81	0.701
C3	Limited resources for delivering CSR programme	4.44	0.910
C4	Unable to decide the best social use of CSR fund	3.52	0.648
C5	Access to finance	3.75	0.690
C6	Company's internal value system or culture	4.63	0.851
C7	Lack of top management commitment	3.23	0.657
C8	Lack of effective strategic planning for CSR	3.75	0.690
C9	Absence of significant benefits for CSR implementation to enterprise	3.12	0.635
C10	Poor supplier and distributor commitment	4.42	0.813
C11	Customers' low willingness to pay for CSR	3.45	0.635
C12	Low involvement of external stakeholders	3.45	0.650
C13	Societal cynicism for CSR as a promotional activity	3.33	0.613
C14	Corruption in system	3.26	0.663
C15	Lack of customer awareness	3.64	0.670
C16	CSR implementation is too complex	2.92	0.347
C17	Absence of organisational reward for employees for doing CSR	2.75	0.327
C18	Lack of external Partner for CSR	2.48	0.295
C19	Lack of benchmarking standards for CSR performances	2.54	0.302

Table 3. Frequency means of CSR barriers

4.2 Structural self-interaction matrix

After identifying the barriers through the questionnaire, the contextual relationships were established in context to CSR implementation problems for pairs of variables. The next step is to develop a structural self-interaction matrix (SSIM) using the following symbols:

- F – if variable "i" will support/help variable "j"; E – if variable "j" will support variable "i"; and
- Y – if the variable "i" and "j" support each other; P -O – if the variable "i" and "j" are not related.

There are 15 barriers identified based on the survey and they are coded in SSIM based on the above criteria. The SSIM of barriers identified is presented in Table 4 below.

4.3 Reachability matrix

The next step is to create an initial SSIM reachability matrix, which is a binary matrix and the transitivity matrix. It is prepared as per the rules put forth by Sushil (2012). Table 5 shows the original matrix of reachability.

4.4 Final reachability matrix

A crucial step in ISM is the transitivity of the contextual interaction. After preparation of the initial reachability matrix, the final reachability matrix is prepared after integrating the transitivity law (Sushil, 2012) as presented in Table 6. The driving force and the dependency power of each vector are determined in the final reachability matrix. For each variable, the driving force is the cumulative number of variables it can help to accomplish (including the variable itself). On the other hand, dependence is the total number of variables that may help accomplish it (including the variable itself). These driving forces and dependencies would later be used under the MICMAC analysis to identify problems into the four classes of autonomous, dependent, linkage and drivers (independent) (Sushil, 2012). The final reachability matrix is provided in the following table, where transitivity was marked as 1.

Table 4.
Structural self-interaction matrix

Barriers	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	F	E	F	P	E	E	F	F	F	O	F	F	P	F
2		E	P	P	E	E	Y	F	P	E	E	E	P	P
3			F	E	E	E	F	F	P	F	F	F	P	F
4				E	E	E	E	Y	E	Y	Y	E	P	E
5					P	F	F	E	P	P	F	P	P	F
6						E	F	F	F	F	F	F	P	F
7							F	Y	Y	Y	F	Y	E	Y
8								F	P	F	F	F	P	F
9									E	E	E	E	P	E
10										F	P	P	P	P
11											E	Y	P	E
12												P	E	P
13													E	E
14														P
15														

Barriers	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
C1	1	0	1	0	0	0	0	1	1	1	0	1	1	0	1
C2	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0
C3	1	1	1	1	0	0	0	1	1	0	1	1	1	0	1
C4	0	0	0	1	0	0	0	0	1	0	1	1	0	0	0
C5	0	0	1	1	1	0	1	1	0	0	0	1	0	0	1
C6	1	1	1	1	0	1	0	1	1	1	1	1	1	0	1
C7	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1
C8	0	1	0	1	0	0	0	1	1	0	1	1	1	0	1
C9	0	0	0	1	1	0	1	0	1	0	0	0	0	0	0
C10	0	0	0	1	0	0	1	0	1	1	1	0	0	0	0
C11	0	1	0	1	0	0	1	0	1	0	1	0	1	0	0
C12	0	1	0	1	0	0	0	0	1	0	1	1	0	0	0
C13	0	1	0	1	0	0	1	0	1	0	1	0	1	0	0
C14	0	0	0	0	0	0	1	0	0	0	0	1	1	1	0
C15	0	0	0	1	0	0	1	0	1	0	1	0	1	0	1

Table 5.
Initial reachability matrix

Barriers	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	Driving power
C1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	13
C2	0	1	0	1	1	0	1	1	1	0	1	1	1	0	1	10
C3	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	13
C4	0	1	0	1	1	0	1	0	1	0	1	1	1	0	0	8
C5	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	14
C6	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	14
C7	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	14
C8	0	1	0	1	1	0	1	1	1	0	1	1	1	0	1	10
C9	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	14
C10	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	14
C11	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	14
C12	0	1	0	1	1	0	1	1	1	0	1	1	1	0	0	9
C13	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	14
C14	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	14
C15	0	0	0	1	0	0	1	0	1	0	1	0	1	0	1	6
Dependence power	10	14	10	15	13	8	15	13	15	10	15	14	15	1	13	181

Table 6.
Final reachability matrix

4.5 Partition of reachability matrix

The next step is level partitioning using iterations after compiling the final reachability matrix. Partitioning aimed to find each variable's hierarchy. The steps in constructing the matrix of reachability are (Sushil, 2012):

- Firstly, for each variable, the reachability (R) and antecedent (A) sets are determined. The reachability set consists of the component itself and the other components that it can help to accomplish. On the other hand, the antecedent set consists of the element itself and the other elements that can assist in achieving it.
- Then, the intersection of these sets i.e. (R∩A) is derived for all the elements.
- The top-level element in the ISM hierarchy is how the reachability and the intersection sets are the same.

- The top-level element of the hierarchy would not help in achieving any other element above their own.
- Once the top-level element is identified, it is separated out from the other aspects.
- The next stage of elements is determined using the same method. The digraph and final model are constructed using these stages.
- Then by the same process, the next level of elements is calculated. These identified levels help in building the digraph and the final model.

4.6 Development of digraph

Based on the level iterations, the elements are arranged graphically and the digraph is developed as per their relationships shown in the final reachability matrix. The indirect links are eliminated (Goyal and Kumar, 2017) by examining their interpretation and only those transitive relationships are retained and shown in the digraph, whose understanding is crucial in Figure 2.

4.7 Interpretive structural model

Finally, after replacing the variable coding with the interpretation of variables in the digraph, the interpretive structural model is developed, as shown in Figure 3 below. From the figure, it can be interpreted that organisational culture is the most crucial barrier that leads to other CSR barriers.

5 Findings of results

5.1 Matrice d'Impacts croises-multiplication appliqué an classment analysis

The 15 barriers are classified by transforming the final reachability matrix into a MICMAC diagram, which evaluates the driving and dependence powers of CSR implementation barriers in medium-sized manufacturing enterprises. The barriers obtained through the final reachability matrix are further grouped into four quadrants.

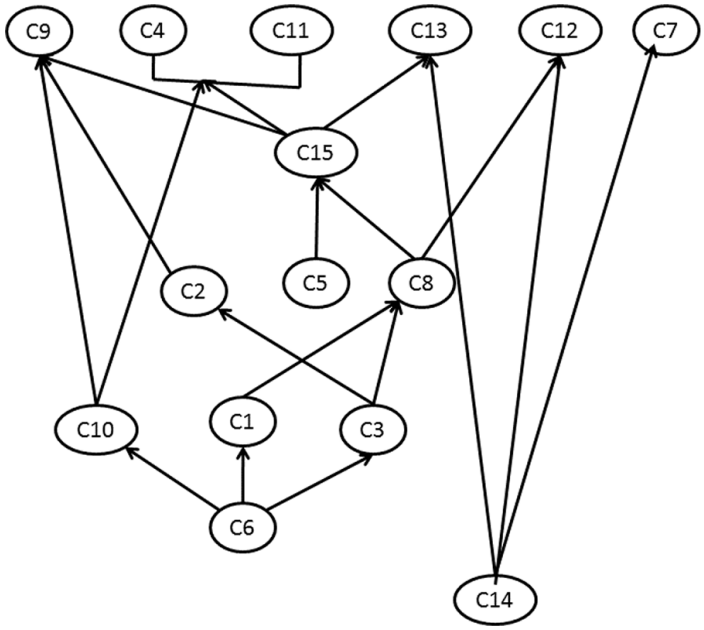
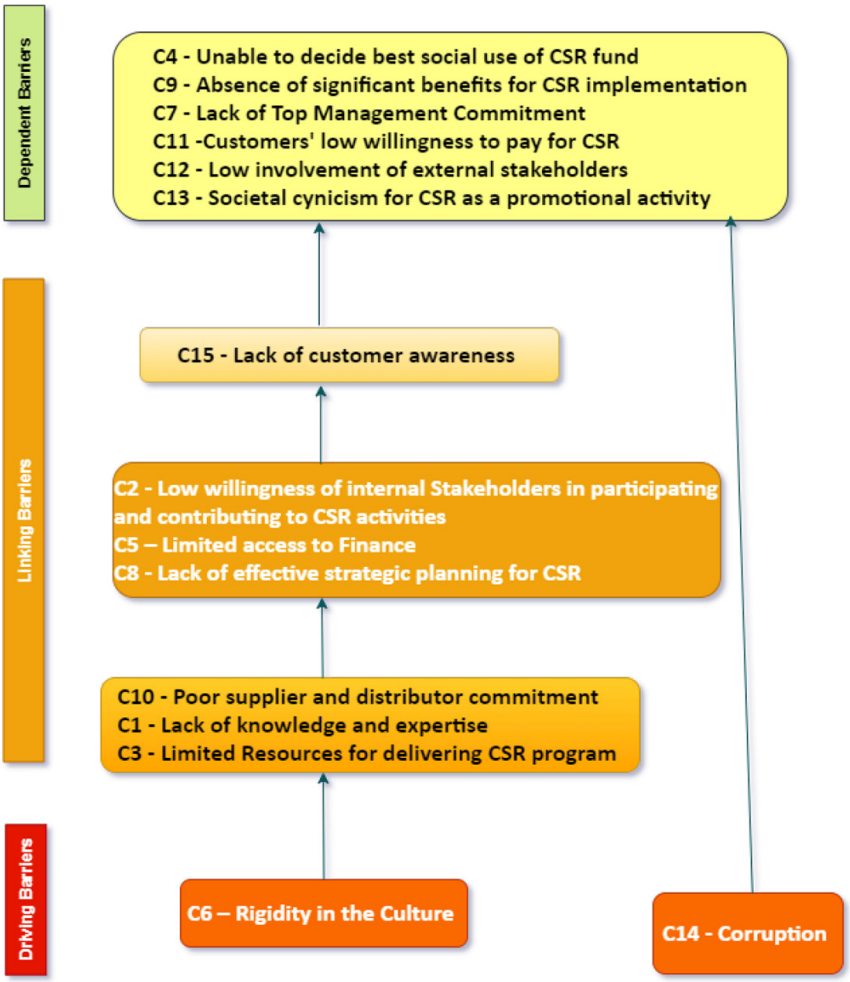


Figure 2.
Diagraph of CSR barriers



Source: Author's own representation

Figure 3.
Categorisation of barriers as per interpretive structural model – final output

The first quadrant of the autonomous variable is both weak drivers and weak dependents, which means that these variables do not influence the other variables of the system, as there are other factors that are dependent on these impediments. The second quadrant of dependent variables includes C4 (Unable to decide the best use of CSR fund) and C15 (Access to Finance) has the highest dependence power and low driving power. Among them, corruption has the highest driving power and lowest dependence, which indicates that it can be treated as the most influential barrier. So, it can be deduced that these variables are highly dependent on other driving variables. The third quadrant is linkage variable has 11 variables i.e. C1 (Lack of knowledge and expertise), C2 (Low willingness of internal stakeholders to participate in CSR activities), C3 (Limited Resources for delivering CSR programme), C5 (Limited access to Finance), C7 (Lack of Top Management Commitment), C8 (Lack of effective strategic planning for CSR), C9 (Absence of significant benefits for CSR implementation to enterprise), C10 (Poor supplier and distributor commitment), C11 (Customers' low willingness to pay for CSR), C12 (Low involvement of external stakeholders) and C13 (Societal cynicism for CSR as promotional activity). Any action directed at these 11 barriers can affect other barriers and have a feedback influence on them.

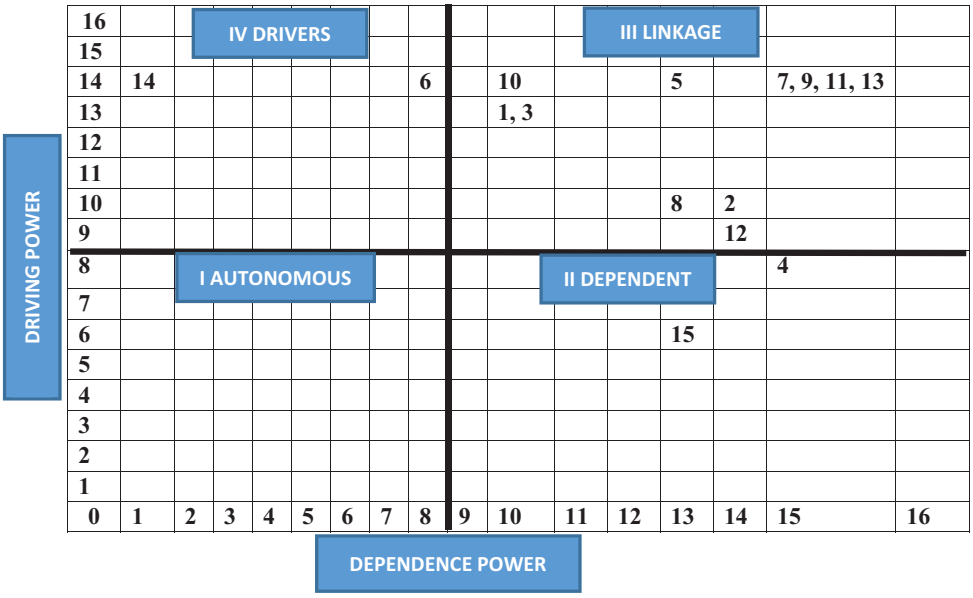


Figure 4.
MICMAC analysis

Fourth quadrant has barriers with high driving force including C6 (Rigidity in culture) and C14 (Corrupt Regulatory Authorities) that possess the highest driving power but the lowest dependence power. Thus, these barriers have the most substantial capability to influence other barriers and should be placed as the highest priority to address. Therefore, medium-sized enterprises need to emphasise these barriers to improve the effectiveness of their CSR activities. The MICMAC analysis of barriers to CSR implementation at medium manufacturing enterprises is presented in Figure 4.

5.2 Discussion of results

The results from the ISM methodology extracted six levels, as shown in Table 7. At iteration level six, Corruption (C14) is considered the most significant barrier to CSR implementation. The perception regarding social responsibility in a developing country is influenced by corruption (Jamali *et al.*, 2017). Corruption is immemorial time and is a significant challenge for modern society, as per World Bank President James Wolfensohn1.

The extensive research on corruption in social sciences and humanities highlight local corruption and corporate corruption influence the enterprise's CSR engagements (Ucar and Staer, 2020; Castro, Phillips and Ansari, 2020). The recurring payment of bribes to government officials to obtain relevant environmental certificates is common in many developing countries (Hasan, 2016; Svensson, 2003; Arevalo and Aravind, 2011). Enterprises operating in a highly corrupt environment earn disinsertion in terms of reduced investment, reduced public expenditure, reduced profit, unproductive behaviour and their CSR engagement is viewed as less trustworthy (Ucar and Staer, 2020). Thus, corruption is a major barrier for CSR implementation, reinforcing weak governance, especially in low-income countries (Demuijnck and Ngnodjom, 2013), which hinders social responsibility performance (Hasan, 2016; Ramasobana and Fatoki, 2014). The output of the current ISM model is unique, as the previous ISM studies did not consider corruption to influence CSR implementation. Our study's results align with previous studies that corruption is the major barrier to CSR implementation (Demuijnck and Ngnodjom, 2013; Hasan, 2016; Ramasobana and Fatoki, 2014; Ucar and Staer, 2020).

Iteration				
Barriers	**Reachability**	**Antecedent**	**Intersection**	**Iteration and level**
4	2, 4, 5, 7, 9, 11, 12, 13	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15	2, 4, 5, 7, 9, 11, 12, 13	I
7	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15	I
9	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15	I
11	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15	I
12	2, 4, 5, 7, 8, 9, 11, 12, 13	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14	2, 4, 5, 7, 8, 9, 11, 12, 13	I
13	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15	I
15	15	1, 2, 3, 5, 6, 8, 10, 14, 15	15	II
2	2, 5, 8	1, 2, 3, 5, 6, 8, 10, 14	2, 5, 8	III
5	1, 2, 3, 5, 6, 8, 10	1, 2, 3, 5, 6, 8, 10	1, 2, 3, 5, 6, 8, 10	III
8	2, 5, 8	1, 2, 3, 5, 6, 8, 10, 14	2, 5, 8	III
1	1, 3, 10	1, 3, 6, 10, 14	1, 3, 10	IV
3	1, 3, 10	1, 3, 6, 10, 14	1, 3, 10	IV
10	1, 3, 6, 10	1, 3, 6, 10, 14	1, 3, 6, 10	IV
6	6	6, 14	6	V
14	14	14	14	VI

Table 7.
Level partition for barriers: Iteration I – iteration VI

Rigidity in culture (C6) is another factor that plays a vital role in CSR adoption among medium-sized manufacturing enterprises in developing countries. Most of the medium-sized enterprises in developing countries believe in the traditional CSR approach that considers social welfare as the governments' role due to their less exposure to the concept (Raynard and Forstater, 2002). The enterprises with such an attitude may have a rigid organisational culture and value system that lacks a pro-active approach to social responsibility and considers it an additional burden (Shen *et al.*, 2015). Our results support the outcomes of ISM models in prior studies (Kumar *et al.*, 2020,2019a, 2019b) supported organisational culture as driving variables, as culture in the organisation supports multi-skilling, competitive thinking and team building for problem-solving.

At level four, the barriers are Poor supplier and distributor commitment (C10), Lack of knowledge and expertise (C1) and Limited resources for delivering CSR programmes (C3). These barriers are essential for medium-sized manufacturing enterprises in implementing CSR. The most important variables in terms of driving power are Lack of knowledge and expertise and Limited Resources for delivering CSR programmes. Lack of resources is also one of the core challenges of CSR implementation in medium-sized manufacturing enterprises (Baumann-Pauly *et al.*, 2013) and it exposes them to various types of operational risk (Faisal *et al.*, 2007). Insufficient CSR knowledge may hinder its adoption and it is one of the major constraints for organisations (Mzembe *et al.*, 2016). Thus, making senior management and employees acquainted with the benefits of CSR may prove crucial for its implementation. Enterprises providing CSR training to employees may benefit them and promote optimism in society. This outcome is complementary with those of Faisal (2010) evidencing the lack of CSR knowledge driving reluctance of supply chain partners in CSR involvement. The results of our study demonstrated both barriers at the same level for medium manufacturing enterprises.

At level three, the lower willingness of internal stakeholders to participate in CSR, ineffective CSR planning and lower access to financial resources fall in the category of linkage variables. These variables are influenced by driving variables and impact other variables in the model. Lower resource allocation to CSR implementation causes the internal stakeholders' lower willingness to participate in CSR (C2). Due to the availability of lower resources, the organisations' priority is to allocate them to their core activities first so that their businesses are profitable. The leftover financial resources are thereafter assigned to social activities (Yuen and Lim, 2016). Thus, lower resource allocation to CSR activities may discourage employees from participating due to a lack of prosocial incentives (Cassar and Meier, 2017). Deficient knowledge of CSR implementation causes ineffective CSR planning (C8). Also, a Lack of CSR knowledge and training fails to connect with the enterprise's proficiency leading to undertaking unrealistic initiatives due to ineffective planning (Rana and Sahay, 2018). Poor planning increases the risk of failure of CSR projects leading to a component of poor corporate valuation. Another major barrier is lower access to financial resources (C5), leading to ineffective CSR planning for CSR implementation. Also, top management believing enterprises do not have sufficient financial resources are less likely to favourably engage in CSR activities. These results support those of Dixit and Priya (2021). If the CSR activities support the competitive advantage of firms, it becomes the strategic decision to increase the competitive advantage and higher return on investment (Faisal, 2010; Husted, 2003).

At level two, Lack of access to financial resources and ineffective CSR planning lead to a lack of customer awareness (15) as a barrier to CSR implementation. Lack of customer awareness has become an issue of CSR implementation in many developing countries (Bux *et al.*, 2020). The burden of increased cost due to CSR initiatives is passed on to customers

(Habel *et al.*, 2016). The customers only pay attention to the product's price and may not be willing to pay the CSR premium (Yuen *et al.*, 2016). The result of the study is supported by the work of Gupta (2011) and Shen *et al.* (2015) evidencing that customer awareness is an important motivation of CSR and their willingness is crucial to its success.

Finally, at level one iteration of ISM, lack of customer awareness leads to their non-willingness to pay a higher price (C11) and CSR is not valued by them. It signifies low bargaining power of enterprises with their customers, absence of significant benefit for CSR implementation to enterprises (C9), unable to decide the best use of CSR funds (C4) and societal cynicism for CSR as promotional activity (C13). Ineffective CSR planning leads to low involvement of effective stakeholders (C12) and corruption in the system decreases the commitment level of top management (C7) towards CSR activities. One possible explanation of this finding can be rigidity in an organisational culture that may not favour CSR. Thus, corruption and a rigid culture and value system are the most critical barrier to CSR implementation in medium-sized manufacturing enterprises.

5.3 Managerial Implications

Globalisation has provided tremendous opportunities to medium-sized manufacturing enterprises and has proved to be a game-changer for them. Many companies have voluntarily incorporated social and responsible practices such as green manufacturing processes, CSR and adopting environmentally-friendly practices to avail of these developmental opportunities. However, CSR practices are not easy in medium-sized enterprises due to their size and barriers across various industries. This study of CSR implementation barriers faced by medium-sized manufacturing enterprises aids policymakers, industrial practitioners and managers in recognise, evaluating and preparing strategies to mitigate its impact. Knowing the interrelationships among CSR implementation barriers will enable the managers to realign their CSR planning. The successful implementation of socially responsible practices may help them achieve economic, social and environmental benefits (Hohnen and Potts, 2007). The current study conducted in the medium manufacturing sector operating in India help in understanding the dynamics of chemical-based, electrical machinery and transportation, machinery and parts expect electrical, rubber plastic and petro-based and metal-based industries. Limited availability of literature of interaction in analysing the interaction among barriers of CSR among medium enterprises is addressed in the study. The research attempts to bridge the gap in the studies exploring CSR barriers in medium manufacturing enterprises. However, the outcome of the current study has various policy implications for CSR implementation. These are below:

- **Addressing barriers at level six**: Corruption in the system is a crucial barrier to CSR implementation in medium enterprises. Corruption in the system affects citizens, private institutions, public institutions and the governments by generating unfair competition, jeopardising economic growth and driving away investments [United Nations Office on Drugs and Crime (UNODC), 2009]. Corruption is incompatible with sustainable development goals owing to its social, economic and environmental damages that lead to reduced profits (Branco and Delgado, 2012). Thus, controlling corruption needs promoting the environment of integrity by promoting best practices, formulating procedures to disseminate information on corruption, creating internal control, creating of anti-corruption policy within the enterprise and a collective change in behaviour by most enterprises is required. The enterprises fighting against corruption and reporting CSR activities thereof make them and others working with them more accountable to the public. Higher

disclosures will improve the understanding of the enterprises of ways to combat corruption and develop an efficient risk management system.

- **Addressing barriers at level five:** The rigid organisational culture is the core barrier to CSR implementation. Therefore, enterprises should plan their CSR projects considering this barrier. To have a CSR-friendly culture, governments and enterprises should facilitate CSR training. The governments should set up benchmark CSR initiatives for enterprises of the same stature. For outstanding performers, benefits can be passed on in terms of subsidies, ease in taxation, funding and relief in international transactions duties. With a positive outlook, medium-sized enterprises should consider CSR as a core activity and pay more attention to its implementation.
- **Addressing barriers at level four:** The barriers at level four poor supplier and distributor commitment (C10), lack of knowledge and expertise (C1) and limited resources for delivering CSR programmes (C3) are critical. Thoughtfully conceptualised, designed and executed CSR training related to CSR guidelines as per company act, 2013 could improve society's understanding of CSR and its benefits. It may result in a favourable attitude of suppliers and distributors towards social responsibility. It could lead to enterprises and their suppliers jointly addressing the CSR issues and their involvement in CSR-related industrial networks. Also, enhanced CSR knowledge may improve CSR projects in the planning phase, leading to optimising resource allocation.
- **Addressing barriers at level three:** The next level consists of the low willingness of internal stakeholders to participate in CSR (C2), ineffective CSR planning (C8) and lower access to financial resources (C5). The most significant of these barriers are lower access to financial resources and the willingness of internal stakeholders. The enterprises with lesser financial resources should collaborate with other institutions or agencies in the cluster. These collaborations can be in vertical perspective i.e. within the same industry or horizontal perspective i.e. outside the industry such as non government organisations (NGOs), industrial associations, enterprises from different industries and governments to address social and environmental issues. Also, appropriate CSR training and allocating substantial resources may lead to effective CSR planning and internal stakeholders can also enjoy social involvement incentives.
- **Addressing barriers at level two:** The barrier at the second level is the lack of customer awareness that can be addressed by providing knowledge about CSR by governments, industries, Universities/educational institutions and non-government organisations. Also, enterprises taking up CSR awareness projects can increase customers' knowledge base and improve social activities performance. Socially-aware customers are more likely to reduce their footprints by demanding environmentally- and socially responsible products and support CSR initiatives.
- **Addressing barriers at level one:** The barriers in the first level are the inability to decide the best use of CSR funds, customer's low interest in CSR and societal cynicism for CSR as promotional activity can be addressed by providing CSR training and creating knowledge awareness regarding social responsibility among the general public. In this regard, crucial steps need to be taken by governments and industry bodies. The increased awareness and knowledge may enable the enterprises to gain significant benefits from their CSR activities and improve top management's commitment towards social responsibility. Providing CSR training

may help enterprises better use the funds in CSR activities. The communication of successful CSR projects may also attract external stakeholders such as NGOs and activists. to get involved in the enterprises' social responsibility activities.

6. Conclusion

This study uses ISM methodology for evaluating barriers to CSR adoption among medium-sized manufacturing enterprises. Manufacturing enterprises have adopted and implemented CSR practices, but they face CSR implementation challenges due to lesser exposure to the concept. This study has attempted to address and prioritise these challenges to identify the most significant barrier to CSR implementation faced by medium-sized manufacturing enterprises. To achieve this objective, a survey was conducted and was later ranked in the order of severity through expert opinion using ISM-MICMAC analysis.

The model results developed in this study concluded that rigidity in the organisation's culture and values and corruption within the organisations and regulatory authorities are the most significant barriers to CSR adoption among medium-sized manufacturing enterprises. Corruption and rigidity in the organisational culture lead to lower CSR knowledge adoption, devoting limited resources to CSR activity and reducing suppliers' and distributors' commitment towards enterprises' responsible behaviour. Inadequate social and environmental knowledge is the major reason for ineffective CSR planning. Devoting lower resources to CSR programmes leads to lower motivation of internal stakeholders to participate in CSR activities. Due to limited financial resources, the enterprises shall be selective in adopting CSR activity and ineffective CSR planning may limit society's awareness, especially among customers, towards social issues. The limited societal understanding of these social issues may lead to a lack of trust in enterprises' CSR activities, leading to societal cynicism as a promotional activity. Due to lower customer awareness, they may be less willing to cooperate with the enterprise for their CSR activities, leading to the customer's lower willingness to be involved in CSR. The ineffective CSR planning may limit the organisational capability to communicate their social achievements to their stakeholders, both internal and external, which may discourage external stakeholders from participating or being involved in CSR activities. Due to limited financial resources and limited customer awareness, the enterprise may not decide on an appropriate CSR activity that it may execute to create value for society. The stakeholders have limited or no interest in enterprises' CSR activities, thus failing to generate significant benefits to the organisations due to their non-involvement. This signifies the need of the organisations to adopt suitable policies and frameworks that follow sustainable practices.

6.1 Limitations and future research directions

There are mainly two limitations of this study. Firstly, the ISM methodology has limitations arising out of managers' selection for deciding the contextual relationships between variables and the bias of managers. Hence, sometimes it may not reflect the true nature of the problem in the real world.

Secondly, this study considered medium-sized manufacturing enterprises for studying barriers, and hence, the usefulness of its conclusions is limited to such enterprises only. The results of the current research should be extended to study other manufacturing sectors. Undertaking similar studies in other contexts shall establish common barriers across different industries and geographical locations.

The current paper presents the conceptual framework that is built based on the analysis of present trends of CSR in medium manufacturing enterprises. The issues

related to the needs of these enterprises have not been addressed in the literature, thus opening the future scope of actionable work. Similar investigations can be undertaken in an industry over a period of time to explore their barriers. To further assess the validation of CSR practice of medium enterprises, the opinions of beneficiaries of their CSR programme may be explored.

Notes

1. Practising social responsibility without the CSR label https://www.theguardian.com/sustainable-business/blog/smes-social-responsibility-without-csr-label
2. http://dcmsme.gov.in/dips/final%20Report%20Ghaziabad.pdf; http://dcmsme.gov.in/dips/DIP%20Report%20Gautam%20Budh%20Nagar.pdf; http://dcmsme.gov.in/dips/har_sonepat.pdf; http://dcmsme.gov.in/dips/2016-17/DIPS%20GURGAON%202015-16.pdf; http://dcmsme.gov.in/dips/DIP%20Report%20%20Faridabad.pdf (accessed on 12 December 2019)
3. Adapted from Preuss and Perschke (2010)

   See: https://openknowledge.worldbank.org/bitstream/handle/10986/7304/343620PAPER0Vo101OFFICIAL0USE0ONLY1.pdf%3Bsequence%3D1

References

Agudo-Valiente, J.M., Garcés-Ayerbe, C. and Salvador-Figueras, M. (2017), "Corporate social responsibility drivers and barriers according to managers' perception; evidence from Spanish firms", *Sustainability*, Vol. 9 No. 10, p. 1821.

Agyemang, O.S. and Ansong, A. (2017), "Corporate social responsibility and firm performance of Ghanaian SMEs: mediating role of access to Capital and firm reputation", *Journal of Global Responsibility*, Vol. 8 No. 1, pp. 47-62, doi: 10.1108/JGR-03-2016-0007.

Alotaibi, A., Edum-Fotwe, F. and Price, A.D. (2019), "Critical barriers to social responsibility implementation within mega-construction projects: the case of the kingdom of Saudi Arabia", *Sustainability*, Vol. 11 No. 6, p. 1755.

Álvarez Jaramillo, J., Zartha Sossa, J.W. and Orozco Mendoza, G.L. (2019), "Barriers to sustainability for small and medium enterprises in the framework of sustainable development – literature review", *Business Strategy and the Environment*, Vol. 28 No. 4, pp. 512-524.

Apaydin, M., Jiang, G.F., Demirbag, M. and Jamali, D. (2020), "The importance of corporate social responsibility strategic fit and times of economic hardship", *British Journal of Management*, Vol. 32 No. 2.

Aragón-Correa, J.A., Hurtado-Torres, N., Sharma, S. and García-Morales, V.J. (2008), "Environmental strategy and performance in small firms: a resource-based perspective", *Journal of Environmental Management*, Vol. 86 No. 1, pp. 88-103.

Arevalo, J.A. and Aravind, D. (2011), "Corporate social responsibility practices in India: approach, drivers, and barriers", *Corporate Governance: The International Journal of Business in Society*, Vol. 11 No. 4, pp. 399-414, doi: 10.1108/14720701111159244.

Bach-Mortensen, A.M., Lange, B.C. and Montgomery, P. (2018), "Barriers and facilitators to implementing evidence-based interventions among third sector organisations: a systematic review", *Implementation Science*, Vol. 13 No. 1, p. 103.

Baden, D.A., Harwood, I.A. and Woodward, D.G. (2009), "The effect of buyer pressure on suppliers in SMEs to demonstrate CSR practices: an added incentive or counter productive?", *European Management Journal*, Vol. 27 No. 6, pp. 429-441.

Baumann-Pauly, D., Wickert, C., Spence, L.J. and Scherer, A.G. (2013), "Organising corporate social responsibility in small and large firms: Size matters", *Journal of Business Ethics*, Vol. 115 No. 4, pp. 693-705.

Bello, F.G. and Kamanga, G. (2020), "Drivers and barriers of corporate social responsibility in the tourism industry: the case of Malawi", *Development Southern Africa*, Vol. 37 No. 2, pp. 181-196, doi: 10.1080/0376835X.2018.1555028.

Belu, C. and Manescu, C. (2013), "Strategic corporate social responsibility and economic performance", *Applied Economics*, Vol. 45 No. 19, pp. 2751-2764.

Bennington, A.J. and Minutolo, M.C. (2013), "Intent to be a socially responsible small or medium-sized enterprise: theory of planned behavior and leaders' actualising", *J. Mgmt. and Sustainability*, Vol. 3, p. 1.

Besch, K. (2005), "Product-service systems for office furniture: barriers and opportunities on the european market", *Journal of Cleaner Production*, Vol. 13 No. 10-11, pp. 1083-1094.

Bevan, E.A.M. and Yung, P. (2015), "Implementation of corporate social responsibility in Australian construction SMEs", *Engineering, Construction and Architectural Management*, Vol. 22 No. 3, pp. 295-311.

Bhardwaj, B.R. (2016), "Role of green policy on sustainable supply chain management: a model for implementing corporate social responsibility (CSR)", *Benchmarking: An International Journal*, Vol. 23 No. 2, pp. 456-468, doi: 10.1108/BIJ-08-2013-0077.

Bhattacharya, C.B., Sen, S. and Korschun, D. (2008), "Using corporate social responsibility to win the war for talent", *MIT Sloan Management Review*, Vol. 49 No. 2.

Branco, M.C. and Delgado, C. (2012), "Business, social responsibility, and corruption", *Journal of Public Affairs*, Vol. 12 No. 4, pp. 357-365.

Bux, H., Zhang, Z. and Ahmad, N. (2020), "Promoting sustainability through corporate social responsibility implementation in the manufacturing industry: an empirical analysis of barriers using the ISM-MICMAC approach", *Corporate Social Responsibility and Environmental Management*, Vol. 27 No. 4, pp. 1729-1748.

Bux, H., Zhang, Z. and Ahmad, N. (2020), "Promoting sustainability through corporate social responsibility implementation in the manufacturing industry: an empirical analysis of barriers using the ISM-MICMAC approach", *Corporate Social Responsibility and Environmental Management*, pp. 1-20.

Campbell, J.L. (2007), "Why would corporations behave in socially responsible ways? An institutional theory of corporate social responsibility", *Academy of Management Review*, Vol. 32 No. 3, pp. 946-967.

Cantor, D.E., Morrow, P.C. and Montabon, F. (2012), "Engagement in environmental behaviors among supply chain management employees: an organisational support theoretical perspective", *Journal of Supply Chain Management*, Vol. 48 No. 3, pp. 33-51.

Carroll, A.B. (1991), "The pyramid of corporate social responsibility: toward the moral management of organizational stakeholders", *Business Horizons*, Vol. 34 No. 4, pp. 39-48.

Cassar, L. and Meier, S. (2017), "Intentions for doing good matter for doing well: the (negative) signaling value of prosocial incentives (no. w24109)", *National Bureau of Economic Research*,

Castka, P., Balzarova, M.A., Bamber, C.J. and Sharp, J.M. (2004), "How can SMEs effectively implement the CSR agenda? A UK case study perspective", *Corporate Social Responsibility and Environmental Management*, Vol. 11 No. 3, pp. 140-149.

Castro, A., Phillips, N. and Ansari, S. (2020), "Corporate corruption: a review and an agenda for future research", *Academy of Management Annals*, Vol. 14 No. 2, pp. 935-968.

Chang, S.J., Van Witteloostuijn, A. and Eden, L. (2010), "From the editors: common method variance in international business research", *Journal of International Business Studies*, Vol. 41 No. 2, pp. 178-184.

Chang, D.S., Chen, S.H., Hsu, C.W. and Hu, A.H. (2015), "Identifying strategic factors of the implantation CSR in the airline industry: the case of Asia-Pacific airlines", *Sustainability*, Vol. 7 No. 6, pp. 7762-7783.

Chen, S. and Bouvain, P. (2009), "Is corporate responsibility converging? A comparison of corporate responsibility reporting in the USA, UK, Australia, and Germany", *Journal of Business Ethics*, Vol. 87 NO. S1, pp. 299-317.

Cheng, B., Ioannou, I. and Serafeim, G. (2014), "Corporate social responsibility and access to finance", *Strategic Management Journal*, Vol. 35 No. 1, pp. 1-23.

Chikara and Sahay (2008), "Evaluation study of medium enterprises in manufacturing sector", Report DIC-MSME.

Chin, M.K., Hambrick, D.C. and Treviño, L.K. (2013), "Political ideologies of CEOs: the influence of executives' values on corporate social responsibility", *Administrative Science Quarterly*, Vol. 58 No. 2, pp. 197-232.

Choi, J.H., Kim, S. and Yang, D.H. (2018), "Small and medium enterprises and the relation between social performance and financial performance: empirical evidence from Korea", *Sustainability*, Vol. 10 No. 6, p. 1816.

Ciliberti, F., Pontrandolfo, P. and Scozzi, B. (2008), "Investigating corporate social responsibility in supply chains: a SME perspective", *Journal of Cleaner Production*, Vol. 16 No. 15, pp. 1579-1588.

Cochran, P.L. and Wood, R.A. (1984), "Corporate social responsibility and financial performance", *Academy of Management Journal*, Vol. 27 No. 1, pp. 42-56.

Coyle, K. (2005), "Environmental literacy in America: what ten years of NEETF/roper research and related studies say about environmental literacy in the US", *National Environmental Education and Training Foundation*,

Coyte, R., Ricceri, F. and Guthrie, J. (2012), "The management of knowledge resources in SMEs: an Australian case study", *Journal of Knowledge Management*, Vol. 16 No. 5, pp. 789-807, doi: 10.1108/13673271211262817.

Curran, J. and Blackburn, R. (2000), *Researching the Small Enterprise*, Sage.

Dearing, A. (2000), *Sustainable Innovation: Drivers and Barriers. Innovation and the Environment*, OECD: Paris, 103-125.

Demuijnck, G. and Ngnodjom, H. (2013), "Responsibility and informal CSR in formal cameroonian SMEs", *Journal of Business Ethics*, Vol. 112 No. 4, pp. 653-665.

Dex, S. and Scheibl, F. (2001), "Flexible and family-friendly working arrangements in UK-based SMEs: business cases", *British Journal of Industrial Relations*, Vol. 39 No. 3, pp. 411-431.

Dixit, A. and Kumar Pandey, A. (2011), "SMEs and economic growth in India: cointegration analysis", *IUP Journal of Financial Economics*, Vol. 9 No. 2.

Dixit, S.K. and Priya, S.S. (2021), "Barriers to corporate social responsibility: an Indian SME perspective", *International Journal of Emerging Markets*, Vol. ahead-of-print No. ahead-of-print, doi: 10.1108/IJOEM-02-2021-0294.

European Commission (2001), *Green Paper: Promoting a European Framework for Corporate Social Responsibility*, Office for Official Publications of the European Communities.

Faisal, M.N. (2010), "Analysing the barriers to corporate social responsibility in supply chains: an interpretive structural modelling approach", *International Journal of Logistics Research and Applications*, Vol. 13 No. 3, pp. 179-195.

Faisal, M.N., Banwet, D.K. and Shankar, R. (2007), "Supply chain risk management in SMEs: analysing the barriers", *International Journal of Management and Enterprise Development*, Vol. 4 No. 5, p. 588, doi: 10.1504/ijmed.2007.013460.

Fassin, Y. (2008), "SMEs and the fallacy of formalising CSR", *Business Ethics: a European Review*, Vol. 17 No. 4, pp. 364-378.

Fitjar, R.D. (2011), "Little big firms? Corporate social responsibility in small businesses that do not compete against big ones", *Business Ethics: A European Review*, Vol. 20 No. 1, pp. 30-44.

Galbreath, J. (2010), "Drivers of corporate social responsibility: the role of formal strategic planning and firm culture", *British Journal of Management*, Vol. 21 No. 2, pp. 511-525.

Gamerschlag, R., Möller, K. and Verbeeten, F. (2011), "Determinants of voluntary CSR disclosure: empirical evidence from Germany", *Review of Managerial Science*, Vol. 5 Nos 2/3, pp. 233-262.

Garay, L. and Font, X. (2012), "Doing good to do well? Corporate social responsibility reasons, practices and impacts in small and medium accommodation enterprises", *International Journal of Hospitality Management*, Vol. 31 No. 2, pp. 329-337, doi: 10.1016/j.ijhm.2011.04.013.

Gautam, R. and Singh, A. (2010), "Corporate social responsibility practices in India: a study of top 500 companies. Global business and management research", *An International Journal*, Vol. 2 No. 1, pp. 41-56.

Ghazilla, R.A.R., Sakundarini, N., Abdul-Rashid, S.H., Ayub, N.S., Olugu, E.U. and Musa, S.N. (2015), "Drivers and barriers analysis for green manufacturing practices in Malaysian SMEs: a preliminary findings", *Procedia Cirp*, Vol. 26, pp. 658-663.

Govidarajan, V., Srivastava, A. and Enache, L. (2019), "Midsize companies are growing, but struggling to earn profits", *Harvard Business Review*, available at: https://hbr.org/2019/11/midsize-companies-are-growing-but-struggling-to-earn-profits (accessed 5 November 2020)

Govindasamy, V. and Suresh, K. (2017), "Exploring approaches to drivers and barriers of corporate social responsibility implementation in academic literature", *In SHS Web of Conferences* (Vol. 33, p. 00021). EDP Sciences.

Goyal, P. and Kumar, D. (2017), "Modeling the CSR barriers in manufacturing industries", *Benchmarking: An International Journal*, Vol. 24 No. 7, pp. 1871-1890, doi: 10.1108/BIJ-09-2015-0088.

Graafland, J. and Smid, H. (2019), "Decoupling among CSR policies, programs, and impacts: an empirical study", *Business and Society*, Vol. 58 No. 2, pp. 231-267.

Greenfield, W.M. (2004), "In the name of corporate social responsibility", *Business Horizons*, Vol. 47 No. 1, pp. 19-28.

Gupta, S. (2011), "Consumer stakeholder view of corporate social responsibility: a comparative analysis from USA and India", *Social Responsibility Journal*, Vol. 7 No. 3, pp. 363-380, doi: 10.1108/17471111111154518.

Gupta, A., Briscoe, F. and Hambrick, D.C. (2017), "Red, blue, and purple firms: organisational political ideology and corporate social responsibility", *Strategic Management Journal*, Vol. 38 No. 5, pp. 1018-1040.

Habel, J., Schons, L.M., Alavi, S. and Wieseke, J. (2016), "Warm glow or extra charge? The ambivalent effect of corporate social responsibility activities on customers' perceived price fairness", *Journal of Marketing*, Vol. 80 No. 1, pp. 84-105.

Hartmann, M. (2011), "Corporate social responsibility in the food sector", *European Review of Agricultural Economics*, Vol. 38 No. 3, pp. 297-324.

Hasan, M.N. (2016), "Measuring and understanding the engagement of Bangladeshi SMEs with sustainable and socially responsible business practices: an ISO 26000 perspective", *Social Responsibility Journal*, Vol. 12 No. 3, pp. 584-610, doi: 10.1108/SRJ.-08-2015-0125

He, G., Zhang, L., Mol, A.P., Wang, T. and Lu, Y. (2014), "Why small and medium chemical companies continue to pose severe environmental risks in rural China", *Environmental Pollution*, Vol. 185, pp. 158-167.

Heugens, P.P.M.A.R. and Oosterhout, V.J. (2008), "Much ado about nothing: a conceptual critique of corporate social responsibility", *The Oxford Handbook of Corporate Social Responsibility*, pp. 197-223.

Hohnen, P. and Potts, J. (2007), *Corporate Social Responsibility: An Implementation Guide for Business*, International Institute for Sustainable Development= Institut international du développement durable.

Hossain, M.M., Alam, M., Hecimovic, A., Alamgir Hossain, M. and Choudhury Lema, A. (2016), "Contributing barriers to corporate social and environmental responsibility practices in a

developing country: a stakeholder perspective", *Sustainability Accounting, Management and Policy Journal*, Vol. 7 No. 2, pp. 319-346, doi: 10.1108/SAMPJ-09-2014-0056.

Hsu, J.L. and Cheng, M.C. (2012), "What prompts small and medium enterprises to engage in corporate social responsibility? A study from Taiwan", *Corporate Social Responsibility and Environmental Management*, Vol. 19 No. 5, pp. 288-305.

Hudson, M., Smart, A. and Bourne, M. (2001), "Theory and practice in SME performance measurement systems", *International Journal of Operations & Production Management*, Vol. 21 No. 8, pp. 1096-1115, doi: 10.1108/EUM0000000005587.

Husted, B.W. (2003), "Governance choices for corporate social responsibility: to contribute, collaborate or internalize?", *Long Range Planning*, Vol. 36 No. 5, pp. 481-498.

Illia, L., Zyglidopoulos, S.C., Romenti, S., Rodríguez-Cánovas, B. and del Valle Brena, A.G. (2013), "Communicating corporate social responsibility to a cynical public", *MIT Sloan Management Review*, Vol. 54 No. 3, p. 2.

Jahdi, K.S. and Acikdilli, G. (2009), "Marketing communications and corporate social responsibility (CSR): marriage of convenience or shotgun wedding?", *Journal of Business Ethics*, Vol. 88 No. 1, pp. 103-113.

Jamali, D. and Abdallah, H. (2015), "Mainstreaming corporate social responsibility at the core of the business school curriculum", In *Handbook of Research on Business Ethics and Corporate Responsibilities*, (pp. 276-300). IGI Global.

Jamali, D. and Karam, C. (2018), "Corporate social responsibility in developing countries as an emerging field of study", *International Journal of Management Reviews*, Vol. 20 No. 1, pp. 32-61.

Jamali, D., Lund-Thomsen, P. and Jeppesen, S. (2017), "SMEs and CSR in developing countries", *Business and Society*, Vol. 56 No. 1, pp. 11-22.

Janssen, C. and Vanhamme, J. (2015), "Theoretical lenses for understanding the CSR–consumer paradox", *Journal of Business Ethics*, Vol. 130 No. 4, pp. 775-787.

Jenkins, H. (2006), "Small business champions for corporate social responsibility", *Journal of Business Ethics*, Vol. 67 No. 3, pp. 241-256.

Jones, M.T. (1996), "Missing the Forest for the trees: a critique of the social responsibility concept and discourse", *Business and Society*, Vol. 35 No. 1, pp. 7-41.

Kamble, S.S., Gunasekaran, A. and Raut, R.D. (2018), "Analysing the implementation barriers of dual cycling in port container terminal using interpretive structural modeling- Indian context", *International Journal of Logistics Research and Applications*, Vol. 22 No. 2, doi: 10.1080/13675567.2018.1492531.

Kechiche, A. and Soparnot, R. (2012), "CSR within SMEs: Literature review", *International Business Research*, Vol. 5 No. 7, p. 97.

Key, S. (1999), "Toward a new theory of the firm: a critique of stakeholder 'theory'", *Management Decision*, Vol. 37 No. 4, pp. 317-328, doi: 10.1108/00251749910269366.

Kumar, D., Goyal, P. and Kumar, V. (2019a), "Modeling and classification of enablers of CSR in Indian firms", *Journal of Modelling in Management*, Vol. 14 No. 2, pp. 456-475, doi: 10.1108/JM2-07-2018-0100.

Kumar, D., Goyal, P. and Kumar, V. (2019b), "Prioritising CSR barriers in the Indian service industry: a fuzzy AHP approach", *Scientific Annals of Economics and Business*, Vol. 66 No. 2, pp. 213-233.

Kumar, N., Mathiyazhagan, K. and Mathivathanan, D. (2020), "Modelling the interrelationship between factors for adoption of sustainable lean manufacturing: a business case from the Indian automobile industry", *International Journal of Sustainable Engineering*, Vol. 13 No. 2, pp. 93-107.

Kumar, D.T., Palaniappan, M., Kannan, D. and Shankar, K.M. (2014), "Analysing the CSR issues behind the supplier selection process using ISM approach", *Resources, Conservation and Recycling*, Vol. 92, pp. 268-278.

Kumar, A., Sah, B., Singh, A.R., Deng, Y., He, X., Kumar, P. and Bansal, R.C. (2017), "A review of multi criteria decision making (MCDM) towards sustainable renewable energy development", *Renewable and Sustainable Energy Reviews*, Vol. 69, pp. 596-609.

Kuo, T.C., Kremer, G.E.O., Phuong, N.T. and Hsu, C.-W. (2016), "Motivations and barriers for corporate social responsibility reporting: Evidence from the airline industry", *Journal of Air Transport Management*, Vol. 57, pp. 184-195, doi: 10.1016/j.jairtraman.2016.08.003.

Kuokkanen, H. and Sun, W. (2016), "Social desirability and cynicism: bridging the attitude-behavior gap in CSR surveys", *Emotions and Organisational Governance (Research on Emotion in Organisations*, Vol. 12, pp. 217-247, doi: 10.1108/S1746-979120160000012008. Emerald Group Publishing Limited,

Kusyk, S.M. and Lozano, J.M. (2007), "SME social performance: a four-cell typology of key drivers and barriers on social issues and their implications for stakeholder theory", *Corporate Governance: The International Journal of Business in Society*, Vol. 7 No. 4, pp. 502-515, doi: 10.1108/14720700710820588.

Lamberti, L. and Noci, G. (2012), "The relationship between CSR and corporate strategy in medium-sized companies: evidence from I taly", *Business Ethics: A European Review*, Vol. 21 No. 4, pp. 402-416.

Laudal, T. (2011), "Drivers and barriers of CSR and the size and internationalisation of firms", *Social Responsibility Journal*, Vol. 7 No. 2, pp. 234-256, doi: 10.1108/17471111111141512.

Law, M.M.S., Hills, P. and Hau, B.C.H. (2017), "Engaging employees in sustainable development–a case study of environmental education and awareness training in Hong Kong", *Business Strategy and the Environment*, Vol. 26 No. 1, pp. 84-97.

Lee, M.H., Mak, A.K. and Pang, A. (2012), "Bridging the gap: an exploratory study of corporate social responsibility among SMEs in Singapore", *Journal of Public Relations Research*, Vol. 24 No. 4, pp. 299-317.

Lepoutre, J. and Heene, A. (2006), "Investigating the impact of firm size on small business social responsibility: a critical review", *Journal of Business Ethics*, Vol. 67 No. 3, pp. 257-273.

López-Gamero, M.D., Claver-Cortés, E. and Molina-Azorín, J.F. (2008), "Complementary resources and capabilities for an ethical and environmental management: a qual/quan study", *Journal of Business Ethics*, Vol. 82 No. 3, pp. 701-732.

Luo, J.M., Huang, G.Q. and Lam, C.F. (2019), "Barriers to the implementation of corporate social responsibility in gaming industry", *Journal of Quality Assurance in Hospitality and Tourism*, Vol. 20 No. 5, pp. 528-551.

Luthra, S., Kumar, V., Kumar, S. and Haleem, A. (2011), "Barriers to implement green supply chain management in automobile industry using interpretive structural modeling technique: an Indian perspective", *Journal of Industrial Engineering and Management (JIEM)*, Vol. 4 No. 2, pp. 231-257.

Lyons, M., Bartlett, J. and McDonald, P. (2016), "Corporate social responsibility in junior and mid-tier resources companies operating in developing nations–beyond the public relations offensive", *Resources Policy*, Vol. 50, pp. 204-213.

Madsen, H. and Ulhøi, J.P. (1996), "Corporate environmental and resource management: overview of results from a survey of the present greening situation in Danish companies", Working Paper No. 4, The Aarhus School of Business, Aarhus.

Malek, J. and Desai, T.N. (2019), "Interpretive structural modelling based analysis of sustainable manufacturing enablers", *Journal of Cleaner Production*, Vol. 238, p. 117996.

Maon, F., Swaen, V. and Lindgreen, A. (2017), "One vision, different paths: an investigation of corporate social responsibility initiatives in Europe", *Journal of Business Ethics*, Vol. 143 No. 2, pp. 405-422.

Mascarenhas, R. (2018), "TCS, Kotak bank among NSE companies not to have fully spent CSR money", *The Economic Times*, available at: https://economictimes.indiatimes.com/markets/stocks/news/tcs-kotak-bank-among-nse-companies-not-to-have-fully-spent-csr-money/articleshow/70563119.cms (accessed 30 October 2020)

Mathiyazhagan, K., Govindan, K., NoorulHaq, A. and Geng, Y. (2013), "An ISM approach for the barrier analysis in implementing green supply chain management", *Journal of Cleaner Production*, Vol. 47, pp. 283-297.

Matten, D. and Moon, J. (2008), "Implicit' and 'explicit' CSR: a conceptual framework for a comparative understanding of corporate social responsibility", *Academy of Management Review*, Vol. 33 No. 2, pp. 404-424.

Metaxas, T. and Tsavdaridou, M. (2010), "Corporate social responsibility in Europe: Denmark, Hungary and Greece", *Journal of Contemporary European Studies*, Vol. 18 No. 1, pp. 25-46.

Min, M., Desmoulins-Lebeault, F. and Esposito, M. (2017), "Should pharmaceutical companies engage in corporate social responsibility?", *Journal of Management Development*, Vol. 36 No. 1, pp. 58-70, doi: 10.1108/JMD-09-2014-0103.

Moktadir, A., Rahman, T., Jabbour, C.J.C., Ali, S.M. and Kabir, G. (2018), "Prioritisation of drivers of corporate social responsibility in the footwear industry in an emerging economy: a fuzzy AHP approach", *Journal of Cleaner Production*, Vol. 201, pp. 369-381.

Montabon, F., Morrow, P.C. and Cantor, D.E. (2016), "Promoting environmental citizenship behaviour", *International Journal of Integrated Supply Management*, Vol. 10 No. 1, pp. 63-88.

Morsing, M. and Perrini, F. (2009), "CSR in SMEs: do SMEs matter for the CSR agenda?", *Business Ethics: A European Review*, Vol. 18 No. 1, pp. 1-6.

Muller, A. and Kolk, A. (2010), "Extrinsic and intrinsic drivers of corporate social performance: evidence from foreign and domestic firms in Mexico", *Journal of Management Studies*, Vol. 47 No. 1, pp. 1-26.

Murillo, D. and Lozano, J.M. (2006), "SMEs and CSR: an approach to CSR in their own words", *Journal of Business Ethics*, Vol. 67 No. 3, pp. 227-240.

Mzembe, A.N., Downs, Y. and Meaton, J. (2016), "Corporate social responsibility in Malawi: Antecedents, issues, practices and future directions", In *Corporate Social Responsibility in Sub-Saharan Africa*, (pp. 3-29). Springer, Cham.

Opricovic, S. and Tzeng, G.H. (2003), "Defuzzification within a multicriteria decision model", *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, Vol. 11 No. 5, pp. 635-652.

Oreg, S. (2003), "Resistance to change: developing an individual differences measure", *Journal of Applied Psychology*, Vol. 88 No. 4, p. 680.

Orlitzky, M., Siegel, D.S. and Waldman, D.A. (2011), "Strategic corporate social responsibility and environmental sustainability", *Business and Society*, Vol. 50 No. 1, pp. 6-27.

Pai, N. (2018), "Opinion | the case against mandatory corporate social responsibility", *The Mint*, available at: https://www.livemint.com/opinion/columns/opinion-a-case-against-punishing-non-compliance-of-csr-guidelines-1564734786880.html (accessed 30 October 2020)

Parviainen, T., Lehikoinen, A., Kuikka, S. and Haapasaari, P. (2018), "How can stakeholders promote environmental and social responsibility in the shipping industry?", *WMU Journal of Maritime Affairs*, Vol. 17 No. 1, pp. 49-70.

Poret, S. (2014), "Corporate-NGO partnerships in CSR activities: why and how?", available at: https://hal.archives-ouvertes.fr/hal-01070474/

Porter, M.E. and Kramer, M.R. (2006), "The link between competitive advantage and corporate social responsibility", *Harvard Business Review*, Vol. 84 No. 12, pp. 78-92.

Prasanna, R.P.I.R., Jayasundara, J.M.S.B., Naradda Gamage, S.K., Ekanayake, E.M.S., Rajapakshe, P.S. K. and Abeyrathne, G.A.K.N.J. (2019), "Sustainability of smes in the competition: a systemic review on technological challenges and sme performance", *Journal of Open Innovation: Technology, Market, and Complexity*, Vol. 5 No. 4, p. 100.

Preuss, L. and Perschke, J. (2010), "Slipstreaming the larger boats: social responsibility in medium-sized businesses", *Journal of Business Ethics*, Vol. 92 No. 4, pp. 531-551.

Puppim de Oliveira, J.A. and Jabbour, C.J.C. (2017), "Environmental management, climate change, CSR, and governance in clusters of small firms in developing countries: toward an integrated analytical framework", *Business and Society*, Vol. 56 No. 1, pp. 130-151.

Rahman, T., Ali, S.M., Moktadir, M.A. and Kusi-Sarpong, S. (2020), "Evaluating barriers to implementing green supply chain management: an example from an emerging economy", *Production Planning & Control*, Vol. 31 No. 8, pp. 673-698.

Ramasobana, A.M. and Fatoki, O. (2014), "An investigation into the business social responsibility of micro enterprises in South Africa", *Mediterranean Journal of Social Sciences*, Vol. 5 No. 3, p. 283.

Rana, S. and Sahay, A. (2018), "The role of law in promoting CSR: the case of Indian maharatna public enterprises", *University of Oslo Faculty of Law Legal Studies Research Paper Series*, pp. 2020-2030.

Raut, R., Narkhede, B.E., Gardas, B.B. and Luong, H.T. (2018), "An ISM approach for the barrier analysis in implementing sustainable practices: the Indian oil and gas sector", *Benchmarking: An International Journal*, Vol. 25 No. 4, pp. 1245-1271, doi: 10.1108/BIJ-05-2016-0073.

Raynard, P. and Forstater, M. (2002), *Corporate Social Responsibility: Implications for Small and Medium Enterprises in Developing Countries*, UNIDO.

Rentizelas, A., de Sousa Jabbour, A.B.L., Al Balushi, A.D. and Tuni, A. (2020), "Social sustainability in the oil and gas industry: institutional pressure and the management of sustainable supply chains", *Annals of Operations Research*, Vol. 290 Nos 1/2, pp. 279-300.

Seidel, M., Seidel, R., Tedford, D., Cross, R., Wait, L. and HäMmerle, E. (2009), "Overcoming barriers to implementing environmentally benign manufacturing practices: strategic tools for SMEs", *Environmental Quality Management*, Vol. 18 No. 3, pp. 37-55.

Shen, L., Govindan, K. and Shankar, M. (2015), "Evaluation of barriers of corporate social responsibility using an analytical hierarchy process under a fuzzy environment – a textile case", *Sustainability*, Vol. 7 No. 3, pp. 3493-3514.

Sherer, S.A. (2003), "Critical success factors for manufacturing networks as perceived by network coordinators", *Journal of Small Business Management*, Vol. 41 No. 4, pp. 325-345.

Shibin, K.T., Gunasekaran, A. and Dubey, R. (2017), "Explaining sustainable supply chain performance using a total interpretive structural modeling approach", *Sustainable Production and Consumption*, Vol. 12, pp. 104-118.

Shirodkar, V., Beddewela, E. and Richter, U.H. (2018), "Firm-level determinants of political CSR in emerging economies: Evidence from India", *Journal of Business Ethics*, Vol. 148 No. 3, pp. 673-688.

Singh, M. and Rathi, R. (2021), "Investigation and modeling of lean six sigma barriers in small and medium-sized industries using hybrid ISM-SEM approach", *International Journal of Lean Six Sigma*, Vol. 12 No. 6, doi: 10.1108/IJLSS-09-2020-0146.

Smith, N.C. (2013), "When it comes to CSR, size matters", *Forbes, available at:* http://www.forbes.com/sites/insead/2013/08/14/when-it-comes-to-csr-sizematters/#468a62271b6f (accessed 16 April 2021).

Solano, S.E., Casado, P.P. and Ureba, S.F. (2016), "Analysis of the dissemination of information on corporate social responsibility in urban public transport companies in Spain", *Accounting Journal*, Vol. 19 No. 2, pp. 195-203.

Spence, L.J. (1999), "Does size matter? The state of the art in small business ethics", *Business Ethics: a European Review*, Vol. 8 No. 3, pp. 163-174.

Stoian, C. and Gilman, M. (2017), "Corporate social responsibility that 'pays': a strategic approach to CSR for SMEs", *Journal of Small Business Management*, Vol. 55 No. 1, pp. 5-31.

Sundar, P. (2017), "Five years after CSR became mandatory, what has it really achieved? [web log post]", available at: https://thewire.in/business/five-years-after-csr-became-mandatory-what-has-it-really-achieved (accessed 30 October 2020)

Sushil, S. (2012), "Interpreting the interpretive structural model", *Global Journal of Flexible Systems Management*, Vol. 13 No. 2, pp. 87-106.

Svensson, J. (2003), "Who must pay bribes and how much? Evidence from a cross section of firms", *The Quarterly Journal of Economics*, Vol. 118 No. 1, pp. 207-230.

Sweeney, L. (2007), "Corporate social responsibility in Ireland: barriers and opportunities experienced by SMEs when undertaking CSR", *Corporate Governance: The International Journal of Business in Society*, Vol. 7 No. 4, pp. 516-523, doi: 10.1108/14720700710820597.

Thakkar, J., Kanda, A. and Deshmukh, S.G. (2008), "Interpretive structural modeling (ISM) of IT-enablers for Indian manufacturing SMEs", *Information Management and Computer Security*, Vol. 16 No. 2, pp. 113-136, doi: 10.1108/09685220810879609.

Turyakira, P.K. (2017), "Small and medium-sized enterprises (SMEs) engagement in corporate social responsibility (CSR) in developing countries: literature review", *African Journal of Business Management*, Vol. 11 No. 18, pp. 464-469.

Ucar, E. and Staer, A. (2020), "Local corruption and corporate social responsibility", *Journal of Business Research*, Vol. 116, pp. 266-282.

United Nations Office on Drugs and Crime (UNODC) (2009), *Business Social Responsibility in Combating Corruption*, OFFICE OF THE COMPTROLLER GENERAL, Brazil, available at: https://www.unodc.org/documents/lpo-brazil/Topics_corruption/Publicacoes/English-Guide-Business-Social-Responsibility.pdf (accessed 17 April 2021)

Villena, V.H. and Gioia, D.A. (2020), "A more sustainable supply chain companies tend to focus on their top-tier suppliers, but the real risks come lower down", *Harvard Business Review*, Vol. 98 No. 2, pp. 84-93.

Vives, A. (2005), "Social and environmental responsibility in small and medium enterprises in Latin America (no. 29538)", *The Journal of Corporate Citizenship*, Vol. 21, pp. 39-50.

Waddock, S.A. and Graves, S.B. (1997), "The corporate social performance–financial performance link", *Strategic Management Journal*, Vol. 18 No. 4, pp. 303-319.

Waddock, S.A., Bodwell, C. and Graves, S.B. (2002), "Responsibility: the new business imperative", *Academy of Management Perspectives*, Vol. 16 No. 2, pp. 132-148.

Wang, Z., Reimsbach, D. and Braam, G. (2018), "Political embeddedness and the diffusion of corporate social responsibility practices in China: a trade-off between financial and CSR performance?", *Journal of Cleaner Production*, Vol. 198, pp. 1185-1197.

Ward, H., Wilson, E. and Zarsky, L. (2007), "CSR and developing countries", *Sustainable Development Innovation Briefs*, Vol. 1, pp. 1-8.

Warfield, J.N. (1973), "Binary matrices in system modeling", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 5, pp. 441-449.

Wilms, W.W., Hardcastle, A.J. and Zell, D.M. (1994), "Cultural transformation at NUMMI", *MIT Sloan Management Review*, Vol. 36 No. 1, p. 99.

Yoon, B. and Chung, Y. (2018), "The effects of corporate social responsibility on firm performance: a stakeholder approach", *Journal of Hospitality and Tourism Management*, Vol. 37, pp. 89-96.

Yuen, K.F. and Lim, J.M. (2016), "Barriers to the implementation of strategic corporate social responsibility in shipping", *The Asian Journal of Shipping and Logistics*, Vol. 32 No. 1, pp. 49-57.

Yuen, K.F., Thai, V.V. and Wong, Y.D. (2016), "Are customers willing to pay for corporate social responsibility? A study of individual-specific mediators", *Total Quality Management & Business Excellence*, Vol. 27 Nos 7/8, pp. 912-926.

Zhang, Q., Oo, B.L. and Lim, B.T.H. (2019), "Drivers, motivations, and barriers to the implementation of corporate social responsibility practices by construction enterprises: a review", *Journal of Cleaner Production*, Vol. 210, pp. 563-584.

Zientara, P. (2017), "Socioemotional wealth and corporate social responsibility: a critical analysis", *Journal of Business Ethics*, Vol. 144 No. 1, pp. 185-199.

Zionts, S. (1979), "MCDM – if not a roman numeral, then what?", *Interfaces*, Vol. 9 No. 4, pp. 94-101.

Further reading

Ackoff, R.L. (1999), *Re-Creating the Corporation: A Design of Organisations for the 21st Century*, Oxford University Press, Oxford.

Appuhami, R. and Tashakor, S. (2017), "The impact of audit committee characteristics on CSR disclosure: an analysis of Australian firms", *Australian Accounting Review*, Vol. 27 No. 4, pp. 400-420.

Beldad, A.D., Seijdel, C.T. and de Jong, M.D. (2019), "Managing corporate social responsibility (CSR) together: the effects of stakeholder participation and third-party organisation (TPO) endorsement on CSR initiative effectiveness", *Corporate Reputation Review*, pp. 1-16.

Dash, S.R., Bhal, K.T. and Udgata, J. (2009), "Organisational culture and environmental responsibility: a study of textile industry in India", *International Journal of Indian Culture and Business Management*, Vol. 2 No. 3, pp. 233-246.

De Neve, G. (2009), "Power, inequality and corporate social responsibility: the politics of ethical compliance in the South Indian garment industry", *Economic and Political Weekly*, pp. 63-71.

Durmusoglu, S.S., Nayir, D.Z., Chaudhuri, M., Chen, J., Joens, I. and Scheuer, S. (2018), "Barriers to firm service innovativeness in emerging economies", *Journal of Services Marketing*, Vol. 32 No. 7, pp. 925-944, doi: 10.1108/JSM-11-2016-0411.

Green, A.O. and Hunton-Clarke, L. (2003), "A typology of stakeholder participation for company environmental decision-making", *Business Strategy and the Environment*, Vol. 12 No. 5, pp. 292-299.

Holtbrügge, D. and Oberhauser, M. (2019), "CSR orientation of future top managers in India", *Journal of Indian Business Research*, Vol. 11 No. 2, pp. 162-178, doi: 10.1108/JIBR-01-2018-0039.

Hsueh, C.F. (2015), "A bilevel programming model for corporate social responsibility collaboration in sustainable supply chain management", *Transportation Research Part E: Logistics and Transportation Review*, Vol. 73, pp. 84-95.

Husted, B.W. and Allen, D.B. (2006), "Corporate social responsibility in the multinational enterprise: strategic and institutional approaches", *Journal of International Business Studies*, Vol. 37 No. 6, pp. 838-849.

Inyang, B.J. (2013), "Defining the role engagement of small and medium-sized enterprises (SMEs) in corporate social responsibility (CSR)", *International Business Research*, Vol. 6 No. 5, p. 123.

Keys, T., Malnight, T.W. and Van Der Graaf, K. (2009), "Making the most of corporate social responsibility", *McKinsey Quarterly*, Vol. 36, pp. 38-44.

Kiesnere, A.L. and Baumgartner, R.J. (2019), "Sustainability management in practice: organisational change for sustainability in smaller large-sized companies in Austria", *Sustainability*, Vol. 11 No. 3, p. 572.

KonSULT (2003), "Barriers to implementation. In decision makers' guidebook", available at: http://www.its.leeds.ac.uk/projects/konsult/public/level1/sec10/index.htm

Lattemann, C., Fetscherin, M., Alon, I., Li, S. and Schneider, A.M. (2009), "CSR communication intensity in Chinese and Indian multinational companies", *Corporate Governance: An International Review*, Vol. 17 No. 4, pp. 426-442.

Lindgreen, A., Maon, F., Reast, J. and Yani-De-Soriano, M. (2012), "Guest editorial: corporate social responsibility in controversial industry sectors", *Journal of Business Ethics*, Vol. 110 No. 4, pp. 393-395.

Narwal, M. and Sharma, T. (2008), "Perceptions of corporate social responsibility in India: an empirical study", *Journal of Knowledge Globalization*, Vol. 1 No. 1, pp. 61-79.

Rama, D., Milano, B.J., Salas, S. and Liu, C.H. (2009), "CSR implementation: Developing the capacity for collective action", *Journal of Business Ethics*, Vol. 85 NO. S2, pp. 463-477.

Samy, M., Odemilin, G. and Bampton, R. (2010), "Corporate social responsibility: a strategy for sustainable business success. An analysis of 20 selected British companies", *Corporate Governance: The International Journal of Business in Society*, Vol. 10 No. 2, pp. 203-217, doi: 10.1108/14720701011035710.

Stekelorum, R., Laguir, I., Courrent, J.M. and Jaegler, A. (2018), "Extending CSR in SMEs' upstream supply chains: a dynamic capabilities perspective", *Supply Chain Forum: An International Journal*, Vol. 19 No. 3, pp. 233-251.Taylor and Francis.

Tamajón, L.G. and Font, X. (2013), "Corporate social responsibility in tourism small and medium enterprises evidence from Europe and Latin America", *Tourism Management Perspectives*, Vol. 7, pp. 38-46.

Velasquez, M. and Hester, P.T. (2013), "An analysis of multi-criteria decision making methods", *International Journal of Operations Research*, Vol. 10 No. 2, pp. 56-66.

Wang, J.J., Jing, Y.Y., Zhang, C.F. and Zhao, J.H. (2009), "Review on multi-criteria decision analysis aid in sustainable energy decision-making", *Renewable and Sustainable Energy Reviews*, Vol. 13 No. 9, pp. 2263-2278.

Warfield, J.N. (1973), "Binary matrices in system modeling", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-3 No. 5, pp. 441-449.

Wilcox, D. (1994), *The Guide to Effective Participation*, Brighton: Partnership.

Corresponding author
Gaurav Dawar can be contacted at: gauravdawar@gmail.com

Appendix

Basis	Large enterprises	Medium enterprises	Small enterprises
Ownership	Ownership by institutional and investors	These enterprises are at the first stage of separation due to the growing use of outside funding	Owners invest the funds through personal borrowing from family and friends, angel investors, bank loans and venture capitalist
Access to capital market	Most funds are raised from capital market	Likely to increase the use of formal capital markets.	No exposure to capital market
Governance	Standard guidelines and corporate governance norms issued by SEBI	Increase in the professionalisation of functions and formalising the organisational structure	Small multi-tasking team, informally establishing systems
Interaction with stakeholders	Interaction with a wide range of stakeholders, formal mechanism for interaction, providing information and feedback	Frequent interaction with external stakeholders to support growth	Interaction with a small number of a stakeholder is generally one to one basis, particularistic and personal
Resource-based view managerial expertise	Professional management with firm-specific knowledge and expertise to attain opportunities	Firm-specific knowledge concentrated among a few key personnel	Owners and managers with general expertise
Organisational structure	Formal hierarchical structure and bureaucratic administration process	Signs of functional differentiation owner and manager likely to retain considerable discretion	Informal and flexible structure
Risk management	Proactive strategic risk management	Focus on the development of practices to control operational risk	Owners are fully involved in operations and limited need for monitoring
CSR activities	The CSR activities of these enterprises are executed through a formal process in line with legal guidelines. The activities are undertaken on a substantial scale and undertaken on a national level.	The CSR activities of these enterprises are mostly undertaken as a part of compliance to law owing to limited resources and they are receptive towards its reporting. The CSR activities undertaken depends on the value system of owners or managers and are restricted to local surroundings	The CSR activities undertaken by these enterprises tend to be economical and mostly voluntary. The activities undertaken are tree plantation, blood donation camps, etc

Table A1.
Nature of enterprises: Difference between small, medium and large enterprises[3]

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)
Pune, India. Oct 29-30, 2021

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON) | 978-1-6654-2503-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/SMARTGENCON51891.2021.9645805

Career Recommender System Using Decision Trees

[1st]Massoud Massoudi
Assistant Professor
Information Systems Department
Parwan University
Charikar, Afghanistan
massoud.massoudi@hotmail.com

[2nd]Siyamoy Ghory
Assistant Professor
Information Technology Department
Parwan University
Charikar, Afghanistan
siyamoyghoury@gmail.com

[3rd]Mahboob Massoudi
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
mahboob.massoudi@gmail.com

***Abstract*—With so many career opportunities present in the computer science field, it becomes pretty tedious for individuals to identify their strengths and goals. This study aims to form a relationship between the experiences of an individual and determine the best suitable career path in the computer science field. Decision trees are used to categorize and organize several disciplines of computer science. According to the data analysis findings, some domains are more likely to be pursued by students since they have comparable academic and professional interests. According to the study results, there is a positive relationship between excellent experiences and career trajectories.**

Keywords: Decision trees; Information gain; Gini; Entropy;

I. INTRODUCTION

In today's world, people have so many options with products, movies, restaurants, and their career areas. With so many different job options available in computer science, it can be difficult for individuals to determine their skill and professional objectives. Sometimes an individual may get very confused about which career area to process through. An individual frequently needs to evaluate skills, interests and certifications to decide their career paths. This paper spreads light on a system that can analyze their strengths and help them pursue their careers. As for an individual looking for guidance in career choices and recruiting candidates, it can help them analyze their strengths and choose which role is the best option for them. This can reduce the workload of recruiters by just entering the details of candidates in the system [1]. Recommendation Systems are an integral aspect of daily life since individuals rely on knowledge to make decisions in their best interests. Recommender systems are a subtype of automated Analysis that is used to anticipate consumer behaviour for things. Although several approaches have been created in the past, the search continues to exist due to its widespread use in various applications that customize advice and cope with information asymmetry. Due to the difficulties posed by these requirements, several methods such as memory-based or model-based are utilized. The recommender system is still in need of improvement to become a better system. A recommendation system is a sophisticated system that alerts users to items that may be of interest to them. Amazon.com, movielens.com, and last.fm are just a few examples. The purpose of this study is to compare the limitations of various approaches and methodologies to make appropriate future suggestions [2]. Earlier attempts of career trajectory are made by using convolutional neural networks architecture to extract the semantic features from the short sentence in the resume [3]. K-Nearest Neighbour method, with the K-Nearest Neighbors method and approach, you may do both regression and classification problems using a single machine learning technique. K-Nearest Neighbors is a statistical approach that evaluates the labels of a specific number of data points to create a prediction model about which class the intended data set belongs to forecast the class that the target data set point belongs. Certain factors arose for predicting student careers based upon students' final exam results and interests [3]. Recommendation systems using collaborative filtering, content-based filtering, matrix factorization and a combination of them are widely known to exist, but recommendation using decision trees needs digging. Decision trees [18] are mainly used for classification problems, so they can be used to classify various career streams and recommend the same. Various factors like their strength, interests, skills, communication skills etc., can be considered and trained upon. Using the knowledge of how decision trees are formed, we can intuitively use them for deciding career paths. The study's decision tree is based on completing a series of questions determining and recommending a position. It analyzes the data and shows it in the form of trees, which are easy to grasp. This alternative form of recommendation can form a correlation matrix and visual tree structure, giving us more insight into the system [1]. Several techniques have been explored to give recommendations, including collaborative filtering, content-based recommendation, and hybrid approaches. Different algorithms and methodologies offer suggestions; nevertheless, collaborative filtering and content-based methods both suffer from the same constraints. Numerous researchers have attempted to circumvent these constraints by merging machine learning and content-based methods to create a hybrid solution that used rating and content information. The recommendation system will always be a source of information for academics [13]. In today's insanely busy environment, recommendation algorithms are becoming extremely important. People are perpetually pressed for time due to the plethora of tasks they must complete within the allotted 24 hours. As a result, suggestion mechanisms are critical since they assist them in making the correct choices without depleting their attentional abilities. The purpose of a recommendation system is to discover material that would excite a person. Additionally, it has various features that enable the creation of personalized lists of useful and entertaining material unique to each user/individual. Recommendation systems are Artificial Intelligence-based algorithms that sift through all available options and provide a personalized list of items that are both exciting and appropriate to a customer [14]. A recommender system is a fully complete software system that makes suggestions using at least one enactment. Additionally, recommender systems include a user interface, a corpus of potential recommendations, and an operator who does own the system.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:02:51 UTC from IEEE Xplore. Restrictions apply.

Additionally, several recommender systems employ two or more views to reaching: CiteULike, a tool for locating and maintaining scholarly references, provides users with two options. Generally, previous research in recommender systems has focused on accuracy, scalability, the new user problem, and suggestion explanations [17]. In the current paper, The purpose of this study is to establish a link between an individual's life experiences and to determine the most appropriate career path in the computer science area for that individual. Decision trees are used to categorize and organize several disciplines of computer science. The remainder of the paper is arranged as follows: Section 2 describes the algorithm, Section 3 demonstrates the approach, and Section 4 ends the article.

II. ALGORITHM

A. Decision Trees

The solution we propose to this problem [19] is a decision tree. Trees have several implications in real life. In its simplest form, a decision tree is a sort of flowchart that depicts a logical path to a conclusion. In terms of data analytics, it is a form of method that classifies data using contingent 'control' expressions. A decision tree begins with a single point and splits in two or more orientations. Each branch presents various alternative outcomes, combining many different and random occurrences before a conclusion is obtained. When seen aesthetically, they have a tree-like appearance...hence the name! Decision trees are precious for machine learning and artificial intelligence because they decompose large amounts of data into smaller, more manageable chunks. They are frequently used for prediction analysis, data categorization, and regression in these domains. Please don't be concerned if this all seems a little abstract—we'll offer some examples below to clarify. To begin, let us examine the various components of a tree structure. One such example is utilized in machine learning for classification and regression. A decision tree is made up of root, intermediate, and leaf nodes. Similarly, the beginning node is referred to as the root node. Decision trees follow a top-down approach, with each node splitting the tree further. The leaf node can't be split further and represents the decision or class. For accurate results, it is expected to have a greater depth of the tree [5].

The splitting of nodes can be done using various methods such as Gini, information gain, variance and chi-square. Commonly used decision trees include *Iterative Dichotomiser 3* (ID3), which is a classification technique that employs an optimization technique to decision tree construction by picking the optimal feature that maximizes Information Gain (IG) or minimizes Entropy (H). *Classification and regression trees* (CART) refers to decision tree techniques used for the classification and regression learning process. Leo Breiman, Jerome Friedman, Richard Olshen, and Charles Stone created the Classification and Regression Tree technique, abbreviated as CART, in 1984. and *C4.5* (an extension of ID3) is a well-known data mining algorithm. C4.5 is a Decision Tree Classifier. C4.5 is a data mining technique for creating decision trees. The C4.5 algorithm is quite effective for generating an actionable conclusion based on an analysis of data [9].

B. Information Gain

The randomness in decision trees is measured using the Entropy, i.e., measuring the lack of order in a system [6]. In information theory and machine learning, Kullback–Leibler divergence is a synonym for information gain. However, when discussing decision trees, mutual information is usually used interchangeably with the Kullback–Leibler divergence of a variable's unitary confidence interval from its conditional probability distribution given the other variable. ID3 and C4.5 are two of many algorithms which use the concept of conditional Entropy and information gain for decision trees [7]. Entropy can be calculated as:

$$E(X) = -\sum_{i=1}^{n} p_i * log_2(p_i) \quad (1)$$

Here, n is the number of classifications in the selected column and p_i is the probability of class I [6]. The information gain is then calculated using from (1)

$$IG(X,Y) = E(X) - \sum_{i=1}^{n} (|X_a|/|X|) * E(X_a)) \quad (2)$$

X_a is the set of rows that has value ***an*** in the *target* column. $|X_a|$and$|X|$ Are the number of rows in their corresponding set [7].

Fig. 1. shows a decision tree generated using information gain for splitting nodes. Decision trees in Fig 1. and Fig. 2. are partially shown up to a depth value of 2.

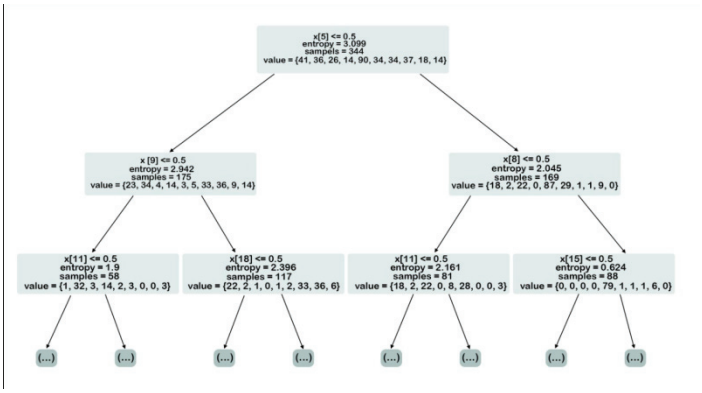


Fig. 1. Decision tree generated via criterion as Entropy

C. Gini Index

The Gini impurity value is one of the approaches used in decision tree algorithms to determine the best first split and future splits from a root node. Gini impurity is another term for the Gini index. What is the Gini coefficient? The Gini index determines the likelihood that a particular characteristic would be miscategorized when picked at random. The probability of classifying a specific variable wrong is evaluated using the Gini measure. CART uses the Gini measure to design binary splits [7]. Gini impurity :

$$GI = 1 - \sum_{i=1}^{n} (p_i)^2 \quad (3)$$

Where p_i is the probability of an object being classified to a particular class. Fig. 2. shows a decision tree generated using the Gini index for splitting nodes.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:02:51 UTC from IEEE Xplore. Restrictions apply.

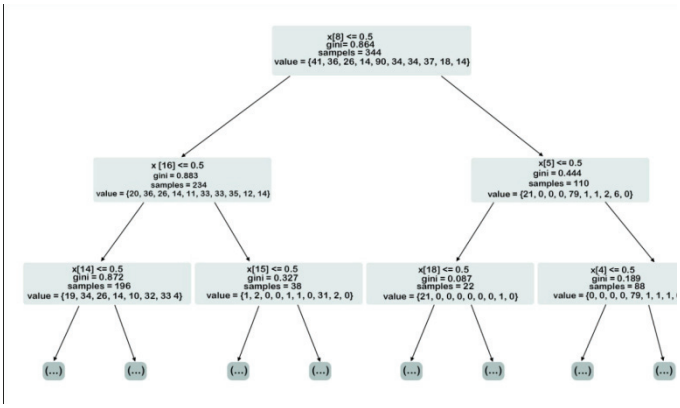


Fig. 2. Decision tree generated using Gini index criterion

III. METHODOLOGY

A practical method for predicting career paths from an individual's characteristics was sought to address this issue. Because aptitude and personality are mostly independent of one another, data from both sources cannot be merged to produce outcomes [9]. The information required should include information on the candidate's talents, credentials, communication skills, and habits; it was preferable if the material was provided in qualitative form. As a result of their comparable academics and training, it was predicted that the bulk of the applicants would fall into one or two classes. Additionally, some of the classes are more popular employment pathways than others. The approach used in this paper is highly suitable because there was a correlation between the features of an individual and his job role. For example, a candidate knowing coding, development skills and interest in technology is expected to become a developer and not a marketing or management agent. So it can be practical to use decision trees in which each node asks these questions and then decide the outcome.

A. Data preprocessing and Analysis

The data that was utilized was obtained from Kaggle. Each row of data includes a series of questions and a work position (label) for the individual answering the questions. There were 524 rows and 20 columns in all, with 18 columns being considered features and the last column (role) serving as the label column. The feature columns answered the candidate's strengths, interests and skills. Label column highlighted which job role is best suited. According to the dataset, the label column had ten classes of functions, and all other feature columns had two types except *communication skills* with three categories. Feature columns represented a set of questions like- *If the individual is good at coding? If they have done any machine learning courses? How are their communication skills? If their interest lies in marketing?*

The dataset was first checked for consistency and redundancy. After checking for duplicate and inconsistent material, the dataset is then reduced to *492* rows and *20* columns. Since all the information was in qualitative form, all the columns were labelled and encoded before processing further. The dataset is then divided into testing and training data in a ratio of *0.30.*

As was predicted, data analysis revealed that the vast majority of the candidates belonged to a specific class, namely, software developers. The bar graph seen in Fig. 3 is an excellent example of this. It contains a variety of applicants for a variety of positions. It indicates that one type of role, namely the developer (dark blue bar), contributes to the bulk of the posts on the list.

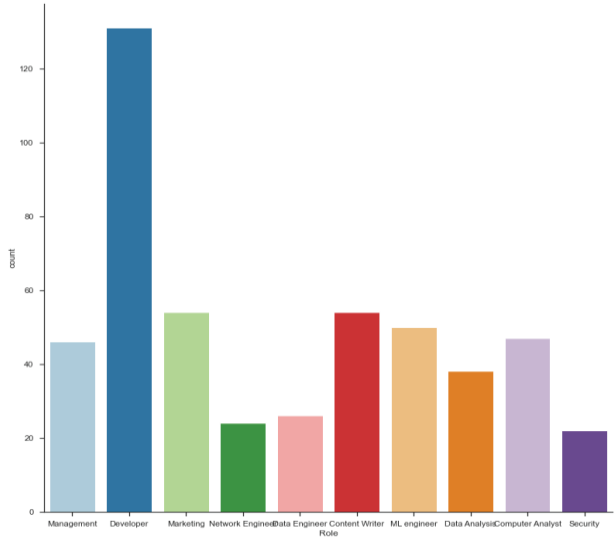


Fig. 3. Number of Candidates versus Role

B. Training and Testing

The dataset was divided into two parts- *feature columns and label columns.*

The decision tree is trained once the data has been divided into two groups: the training set and the test set. The Decision Tree Classifier from the sklearn package is used to train and test the decision tree model. The data is fitted and tested for the methods Gini and Entropy. From a maximum depth of 1 to 10, confusion matrix and accuracy are evaluated for both Gini and entropy criteria. From all the accuracy assessed, the best accuracy is considered, and confusion matrix and decision trees are visualized.

Further analyses of data are done using support vector machine (SVM) classification, and the confusion matrix is visualized. During the training phase, Fig. 1. and Fig. 2. were generated, which gave an in-depth visualization of Gini based and Entropy-based decision trees and their associated variables, respectively. Figure 4 marks the comparison of accuracy graphically between SVM, Information gain and Gini index method.

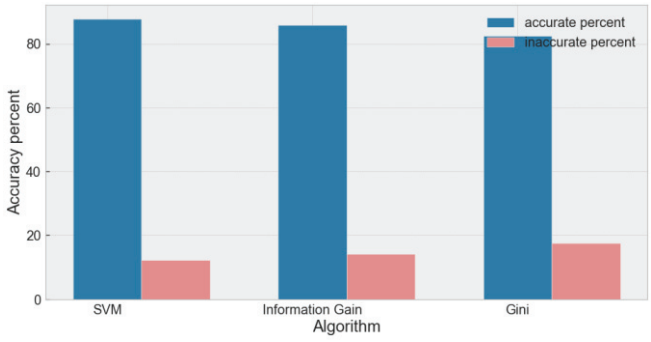


Fig. 4. Accuracy graph.

IV. CONCLUSION

When it comes to categorization jobs, decision trees have shown to be quite helpful. The accuracy of decision trees based on the Gini criteria was found to be 83.10 per cent.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:02:51 UTC from IEEE Xplore. Restrictions apply.

According to the results, the accuracy of decision trees based on the information acquisition technique was 85.81 per cent. Support vector machine gave the highest accuracy of 87.83. It is concluded that decision trees based on information gain, having a tree depth of 6, gave better results than the Gini measure method.

REFERENCES

[1] Sripath Roy, K., Roopkanth, K., Uday Teja, V., Bhavana, V., & Priyanka, J. (2018). Student Career Prediction Using Advanced Machine Learning Techniques. International Journal of Engineering & Technology,7(2.20),26-29.doi:http://dx.doi.org/10.14419/ijet.v7i2.20.11738

[2] B. Bhatt, P. J. Patel, H. Gaudani, and A. Professor, "A Review Paper on Machine Learning-Based Recommendation System," Int. J. Eng. Dev. Res., vol. 2, no. 4, pp. 2321–9939, 2014, [Online]. Available: www.ijedr.org.

[3] He, M., Shen, D., Zhu, Y., He, R., Wang, T., & Zhang, Z. (2019). Career Trajectory Prediction based on CNN. 2019 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI).

[4] Nunsina, Tulus, & Situmorang, Z. (2020). Analysis Optimization K-Nearest Neighbor Algorithm with Certainty Factor in Determining Student Career. 2020 3rd International Conference on Mechanical, Electronics, Computer, and Industrial Technology (MECnIT).

[5] Shankhdhar, A., Agrawal, A., Sharma, D., Chaturvedi, S., & Pushkarna, M. (2020). Intelligent Decision Support System Using Decision Tree Method for Student Career. 2020 International Conference on Power Electronics & IoT Applications in Renewable Energy and Its Control (PARC).

[6] Huaining Sun and Xuegang Hu, "An improved learning algorithm of decision tree based on entropy uncertainty deviation," 2012 IEEE 14th International Conference on Communication Technology, Chengdu, 2012, pp. 799-803, doi: 10.1109/ICCT.2012.6511313.

[7] Navada, A., Ansari, A. N., Patil, S., & Sonkamble, B. A. (2011). Overview of use of decision tree algorithms in machine learning. 2011 IEEE Control and System Graduate Research Colloquium.

[8] Sivagama Sundhari, S. (2011). A knowledge discovery using decision tree by Gini coefficient. 2011 International Conference on Business, Engineering and Industrial Applications.

[9] Rangnekar, R. H., Suratwala, K. P., Krishna, S., & Dhage, S. (2018). Career Prediction Model Using Data Mining and Linear Classification. 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA).

[10] Ade, R., & Deshmukh, P. R. (2014). An incremental ensemble of classifiers as a technique for prediction of student's career choice. 2014 First International Conference on Networks & Soft Computing (ICNSC2014).

[11] Elayidom, S., Idikkula, S. M., Alexander, J., & Ojha, A. (2009). Applying Data Mining Techniques for Placement Chance Prediction. 2009 International Conference on Advances in Computing, Control, and Telecommunication Technologies.

[12] Devasia, T., Vinushree T P, & Hegde, V. (2016). Prediction of students' performance using Educational Data Mining. 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE).

[13] Meenakshi Sharma and Sandeep Mann, "A Survey of Recommender Systems: Approaches and Limitations",ICAECE- 2013.

[14] Supianto, A. A., Julisar Dwitama, A., & Hafis, M. (2018). Decision Tree Usage for Student Graduation Classification: A Comparative Case Study in Faculty of Computer Science Brawijaya University. 2018 International Conference on Sustainable Information Engineering and Technology (SIET).

[15] Rutvija Pandya Jayati Pandya , "C5.0 Algorithm to Improved Decision Tree with Feature Selection and Reduced Error Pruning", International Journal of Computer Applications (0975 – 8887) Volume 117 – No. 16, May 2015.

[16] Huang Ming, Niu Wenying, & Liang Xu. (2009). An improved Decision Tree classification algorithm based on ID3 and the application in score analysis. 2009 Chinese Control and Decision Conference.

[17] C. Davidsson and S. Moritz, "Utilizing implicit feedback and context to recommend mobile applications from first use," ACM Int. Conf. Proceeding Ser., pp. 19–22, 2011, doi: 10.1145/1961634.1961639

[18] M. Massoudi, N. K. Jain and P. Bansal, "Software Defect Prediction using Dimensionality Reduction and Deep Learning," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 2021, pp. 884-893, doi: 10.1109/ICICV50876.2021.9388622.

[19] M. Massoudi, S. Verma and R. Jain, "Urban Sound Classification using CNN," 2021 6th International Conference on Inventive Computation Technologies (ICICT), 2021, pp. 583-589, doi: 10.1109/ICICT50816.2021.9358621.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:02:51 UTC from IEEE Xplore. Restrictions apply.

Clickbait in YouTube: Prevention, Detection and Analysis of the Bait using Ensemble Learning

Peya Mowar*

Department Information Technology, Delhi Technological University, peyajm29@gmail.com

Mini Jain*

Department Information Technology, Delhi Technological University, mini.j2799@gmail.com

Ruchika Goel

Department Information Technology, Delhi Technological University, ruchika2ar@gmail.com

Dinesh Kumar Vishwakarma

Department Information Technology, Delhi Technological University, dinesh@dtu.ac.in

Unscrupulous content creators on YouTube employ deceptive techniques such as spam and clickbait to reach a broad audience and trick users into clicking on their videos to increase their advertisement revenue. Clickbait detection on YouTube requires an in-depth examination and analysis of the intricate relationship between the video content and video descriptors – title and thumbnail. However, the current solutions are mostly centred around the study of video descriptors and other metadata such as likes, tags, comments, etc., and fail to utilize the video content, i.e., both video and audio. Therefore, we introduce a novel model to detect clickbaits on YouTube that consider the relationship between video content and title/thumbnail. The proposed model consists of a stacking classifier framework composed of six base models (K-Nearest Neighbours, Support Vector Machine, XGBoost, Naive Bayes, Logistic Regression, and Multilayer Perceptron) and a meta-classifier (Random Forest). The developed clickbait detection model achieved a high accuracy of 92.89% for the novel BollyBAIT dataset and 95.38% for Misleading Video Dataset (MVD). Additionally, the stated classifier does not use meta-features or other statistics dependent on user interaction with the video (e.g., the number of likes, followers, or comments) for classification, and thus, can be used to detect potential clickbait videos before they are uploaded, thereby preventing the nuisance of clickbaits altogether and improving the users' streaming experience.

CCS CONCEPTS • Information systems ~ World Wide Web ~ Web applications ~ Social Networks • Computing methodologies ~ Artificial Intelligence ~ Computer Vision • Social and professional topics ~ Computing / technology policy ~ Censorship ~ Technology and censorship • Computing methodologies ~ Machine learning ~ Machine learning algorithms ~ Ensemble methods

Additional Keywords and Phrases: clickbait prevention, clickbait detection, ensemble learning, bollywood, youtube

1 INTRODUCTION

The advent and exponential rise of social media, like Twitter, Instagram and YouTube provided their users the opportunity to both create and consume a variety of online content to stay up-to-date with current world events, share their views/knowledge, market their business products etc. Unfortunately, unscrupulous content creators misuse these platforms to publish misleading, poor-quality content which often disseminates false

* Peya Mowar and Mini Jain have contributed to this work equally as first authors.

information[11][21] by employing techniques such as hoaxes[42][46] rumours[24][23][34][45], fake news [25][15][32][43] and clickbaits[42][7]. While the motivations for publishing hoaxes and rumours range from scaremongering to manipulation of popular public opinion, the primary incentive behind posting clickbait content is to gain more clicks on either the content itself (e.g., video) or the link attached to the content (e.g., news headline with link to full article). An increased number of clicks leads to an increase in advertisement revenue for the creator but degrades the quality of user experience and often leaves users feeling unsatisfied, tricked, and annoyed. In the context of YouTube, examples of such content include videos with morphed eye-grabbing thumbnails, exaggerated headlines that include curiosity-inciting phrases such as “You Won’t Believe”, “X Reasons Why” and often fail to deliver on the promised sensational content, or the videos whose content is completely irrelevant or significantly digresses from that suggested by the video descriptors.

Clickbait as a concept was originally found in news media, where captivating headlines had a significant dissonance from the actual content of the article [5][6] .With the increasing popularity of social media, these text-centric baits were additionally seen in platforms such as Twitter and Reddit. As a consequence, researchers were prompted to develop clickbait detection models for identifying clickbaits in text-centric social media [1][44][28]. Subsequently, there has been a shift from text-centric baits, to visual-centric and video-centric baits, wherein content creators make use of enticing thumbnails and videos to grab a user’s attention respectively. This motivated researchers to redefine clickbaits to encompass their different types and develop models to detect them on various other types of social media such as Instagram [12] and YouTube [41][31][38]. Previously, we explored visual-centric baits and created a cross-platform clickbait detection model comprising of a stacking framework architecture [17]. Encouraged by the outcome, we decided to foray into detection of video-centric clickbaits.

Since, clickbait detection and prevention in videos is a largely unexplored field, the primary aim of the conducted research is to develop the Clickbait Prevention and Detection Model (CPDM) for YouTube videos based on only the (i) Video Content, i.e., both the video and associated audio, (ii) Video Title, and (iii) Thumbnail. The video title and thumbnail are used by the viewer to decide to click on a video and, ideally, should be an accurate representation of the video content. The model will determine if they have been exaggerated, manipulated, and engineered to entice the users into clicking on them or are ‘clickbaity.’ Besides, we aim to utilize the information and features extracted from the descriptors and the video to build a model that can determine if the latter fulfils the promise of content suggested by the former. This contrasts with the previously developed models for clickbait detection on YouTube which are heavily dependent on meta-features such as comments, number of likes etc. for the video accumulated over a considerable period of time.

Therefore, to the best of our knowledge, CPDM will be the first model that not only takes into account the intricate relationship between the video contents and its title/thumbnail utilizes it to classify a video as clickbait, but is also completely independent of any meta-features or other features dependent on user engagement with the video. Hence, it can be used for real-time detection of clickbait videos and be used to discard them before they reach a broad audience and lead to poor user experience. Hence, while previous research [41][31][38] focused on developing clickbait detection models, i.e., ‘detect’ them after they have been published and garnered significant user engagement, the proposed model can be used to both ‘prevent’ the publishing of clickbait content and deterioration of content quality and ‘detect’ it if it has already been published.

In our research, we have defined the videos on YouTube as clickbait if they fall into one or more of the following five categories:

1. Misleading content: These contain false information or fake news in either the video descriptors, i.e., headline and thumbnail and/or video content, e.g., A video showing morphed images of an actress's baby bump when she is not pregnant.
2. Spam Content: These videos use tricks such as referring to many famous figures in title and thumbnail that are irrelevant to the content to mislead them into clicking on videos to increase ad revenue or redirect them to malicious websites, e.g., A video with title 'Sultan Full Movie | Salman Khan | Aamir Khan | Katrina Kaif' but the video redirects to a link to watch the full movie instead.
3. False Promises: These videos fail to deliver on the content indicated by the headline or thumbnail, e.g., A video that claims to be the official trailer of an upcoming movie but instead shows press meetings of the lead actors.
4. Exaggerated Headline/Thumbnail: These videos contain exaggerated descriptors that are often embellished with emojis, capitalized words, sentimental and catchy phrases, arrows to manipulate the user into clicking on them, e.g., A video with headline, '5 Extremely Bizarre Indian Movie Scenes That Will Surely Leave You SPEECHLESS | Watch Alone'.
5. Exploits Curiosity Gap: These videos contain descriptors that reveal just enough information to incite the user's curiosity but not enough to satisfy it unless they click on the video, e.g., A video with the headline 'This World-Famous Hollywood Actor got Secretly Married Yesterday | Find out WHY and WHERE' and with a video thumbnail that contains a person's blurred image to arouse the user's interest further.

Examples of YouTube videos for each category are shown in Figure 1.

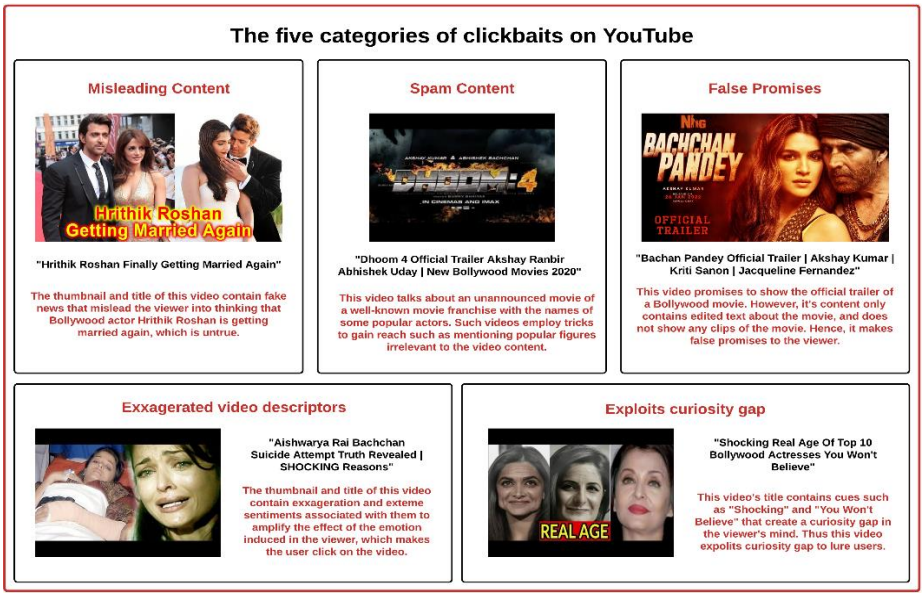


Figure 1: The five categories of clickbaits on YouTube.

The major contributions of our research include:

- Clickbait ‘Prevention’ and ‘Detection’: CPDM uses only the video’s keyframes, its audio, and associated title and thumbnail to classify a video as clickbait or not and is independent of any meta-features or user-dependent features of the video. Since, all features used by CPDM are available at the time of video upload, our model can not only detect existing clickbaits on YouTube but can also be used to identify a clickbait video in real-time before it is made available to public, and thus prevent the problem of clickbait altogether.
- Multilingual Support: CPDM supports identification of clickbait videos in over 200 languages (e.g., English, Hindi, Telugu etc.) and thus, has a larger reach than previously proposed models which can mostly detect clickbait videos in English.
- BollyBAIT Dataset: A novel manually annotated dataset that contains 1000 videos belonging to clickbait and non-clickbait categories. Clickbait videos have been further classified into one or more of the types of clickbait defined above. The videos have been selected from matters related to Bollywood because it is India’s leading entertainment industry with a worldwide audience and reach [47] and hence, draws many clickbait creators and spammers. Also, it is multilingual and contains videos in English, Hindi, or a mixture of both.
- Redefinition of Clickbaits on YouTube: Since, social media platforms such as YouTube are ever-evolving, in terms of technology, available content, its creators and consumers, the accompanying forms of clickbaits are also rapidly expanding. Consequently, even though researchers in the past have attempted to define clickbaits [41][38], these definitions are outdated and not broad enough to encapsulate all types of clickbaits currently present on YouTube. Therefore, we performed a holistic study of clickbaits on YouTube and its previous definitions, to identify its various new and emerging categories and then redefined clickbaits to include a vast majority of them. The updated and broader definition of clickbaits was ultimately incorporated in the proposed BollyBAIT dataset and model, which to the best of our knowledge, can now detect and help in preventing a wide majority of currently prevalent and emerging clickbait techniques.

This paper is organized as follows. Section 2 reviews the related work on defining clickbait for YouTube, various approaches to clickbait detection, research gaps, and future scope. Section 3 introduces the novel clickbait dataset, BollyBAIT, and discusses its features and applications. Section 4 discusses the methodology for feature extraction from video and its descriptors and presents an overview of the architecture of our proposed model, CPDM. Section 5 discusses the results of clickbait identification on BollyBAIT and MVD Dataset using the proposed model. Finally, Section 6 presents the conclusion and explains the future scope for the conducted research.

2 LITERATURE REVIEW

Much of the previous work in clickbait YouTube detection has focused on detecting text-centric clickbaits in articles. However, with the emergence of visual-centric social media and video-streaming platforms like YouTube and Vimeo, newer forms of clickbaits in terms of images and videos surfaced. As a result, researchers redefined clickbaits and proposed models to cater to these newer modalities. This section presents an overview of past definitions of clickbaits and developed schemes to detect them on YouTube.

2.1 Defining Clickbait for YouTube

Due to the ever-evolving nature of social media, the definition of clickbait has evolved too as the intended user-base, use-cases, and the supported content modalities of these platforms diversified. While the early interpretations of clickbaits focused on tabloid journalism, where clickbaits referred to short messages that enticed users to click on the attached link [18], this description was narrow to encompass the variety of clickbaits that came into existence as visual-centric social media became popular. The posts on these contained more descriptors than short textual messages in images, teaser videos, and audios. Therefore, the need for a broader definition of clickbaits that took into consideration the new modalities arose.

In the case of Instagram, clickbaits have been defined as incongruous pairs of text and accompanying images that have been purposely architected to lure users into viewing advertisements and other content to increase their engagement and revenue [12]. For YouTube, researchers have employed various definitions of clickbaits. Zannetteou et al. [41] have defined clickbaits as videos that fall into one or more of the following categories: (i) using dramatic, enticing, and content-irrelevant thumbnails (ii) embedding fake news or false information into the title, thumbnail, or video, (iii) using exaggerated headlines to incite user curiosity and lure them into clicking. However, Shang et al. [31] have defined clickbaits as videos whose content sharply differs from the one promised by the title or thumbnail. Varshney et al. [38] have based their categorization of clickbait on the one proposed by George Loewenstein and defined them as videos that exploit the information gap theory to entice a user into clicking on a video that does not contain the suggested content and eventually degrades the user experience.

2.2 Detecting Clickbaits on YouTube

There have been primarily two approaches to clickbait detection in YouTube, one that took into consideration the content of the video and the other that relied only on video descriptors like title, thumbnail, and other metadata like comments, the number of views, and likes to detect clickbaits. Below, we discuss the past work done in terms of both these approaches.

2.2.1 Content-Agnostic Approaches for Clickbait Detection

Zannetteou et al. [41] developed a deep-learning-based model based on variational autoencoders that do not directly analyse the video content but uses features extracted from data and video descriptors such as headline, tags, thumbnail, comments, number of views, likes, etc. which were collected using the YouTube's Data API for 206k videos. The extracted features were majorly text-based or statistics except for thumbnail image. On the other hand, Shang et al. [31] have adopted a fully content-agnostic approach and only used comments and metadata (e.g., number of views) while building the Online Video Clickbait Protector Scheme (OVCP). Besides not using video content, they did not utilize the descriptors available to users before clicking on the video, such as image and thumbnail, unlike [41].

2.2.2 Methods Utilising Video Content for Clickbait Detection

While most research has focused on developing clickbait detection models utilizing video descriptors and meta-features, only a few recent approaches have used features extracted from the video content, i.e., audio and video, to classify clickbaits. Varshney et al.[38] have incorporated the cosine similarity between the transcript of the video's audio and its title as a feature to inspect if the video delivers on the content as suggested

by the title. However, researchers have not employed any features based on the visual aspect of the video in training their models previously.

2.3 Research Gaps

In [41], Zannetteou et al. have concluded that the YouTube Recommendation System fails to consider clickbaits while recommending related videos and is likely to recommend more clickbait videos when users click on one such video. Therefore, this lacking feature of the YouTube Recommendation System pushes the user into a web of clickbaits, wasting their time and causing user dissatisfaction. Furthermore, they examined two statistics for both clickbait and non-clickbait videos: (i) Number of Unavailable Videos (NUV): Number of videos removed by YouTube or video publisher (ii) Mean of Time Elapsed (MTE): The mean of time elapsed from video publication date till January 10, 2017, for all videos in each category. After performing this analysis, they deduced that YouTube employs no timely corrective measures to remove clickbait videos because of the small percentage of NUV for clickbaits and a high value for MTE.

Accordingly, a model that can prevent clickbaits altogether, i.e., label a video as clickbait before uploading and discoverable by the public, can help tackle the menace of clickbaits on YouTube and improve user experience immensely. This can be achieved by utilizing only the modalities available during upload, like video content (both audio and video), its title, and thumbnail, and not relying on the meta-data generated due to user interaction with video like the number of likes, dislikes, comments, etc. Furthermore, authors in [31] draw attention to their analysis and subsequent conclusion that focus and utilization of only one of the titles, thumbnail [37][40], or the video content [27] cannot be a definitive indicator of the video being a clickbait. Consequently, there is scope to design a model that explores their complex relationship between video content and title/thumbnail.

3 BOLLYBAIT: A BOLLYWOOD-CENTRIC CLICKBAIT DATASET FOR YOUTUBE

We present a novel clickbait dataset with 1000 videos for YouTube titled ‘BollyBAIT’ containing 500 clickbait and 500 real/non-clickbait videos. All videos in the dataset are related to Bollywood and belong to one or more of various categories such as video interviews of celebrities, trailer launch videos, songs, trailers of Bollywood movies, and news and updates related to ongoing happenings in the industry or Bollywood celebrities’ lives. The primary reason for focusing on Bollywood-related videos on YouTube is the popularity of the Bollywood industry. The Indian film industry, especially Bollywood, attracts a lot of attention [47] with a worldwide audience [48]. Therefore, many unscrupulous content creators who want to draw mileage from the popularity of Bollywood to gain advertisement revenue create spam and clickbait videos to reach a vast audience, resulting in an enormous number of clickbait videos available on YouTube.

The videos have been manually collected by searching for queries such as "Upcoming Bollywood movie trailers", "Anushka Sharma pregnancy", "Full movie Bollywood", "Latest news Shah Rukh", "Shocking news Kareena Kapoor". The queries were formed by including famous Bollywood celebrities like Anushka Sharma, Kareena Kapoor, Shah Rukh Khan, and phrases of interest like pregnancy, latest movies, shocking news, etc. After collecting the videos, they were manually labelled by three reviewers as clickbait or real by analysing the (i) Title, (ii) Thumbnail, and (iii) Video content. For each video, it was examined if they belonged to the categories that were defined earlier in Section I of the paper. A video was labelled as clickbait if it fell into one or more of

these categories; otherwise, it was labelled as real. Our dataset covers a massive number of clickbait videos related to Bollywood as shown in Figure 2.

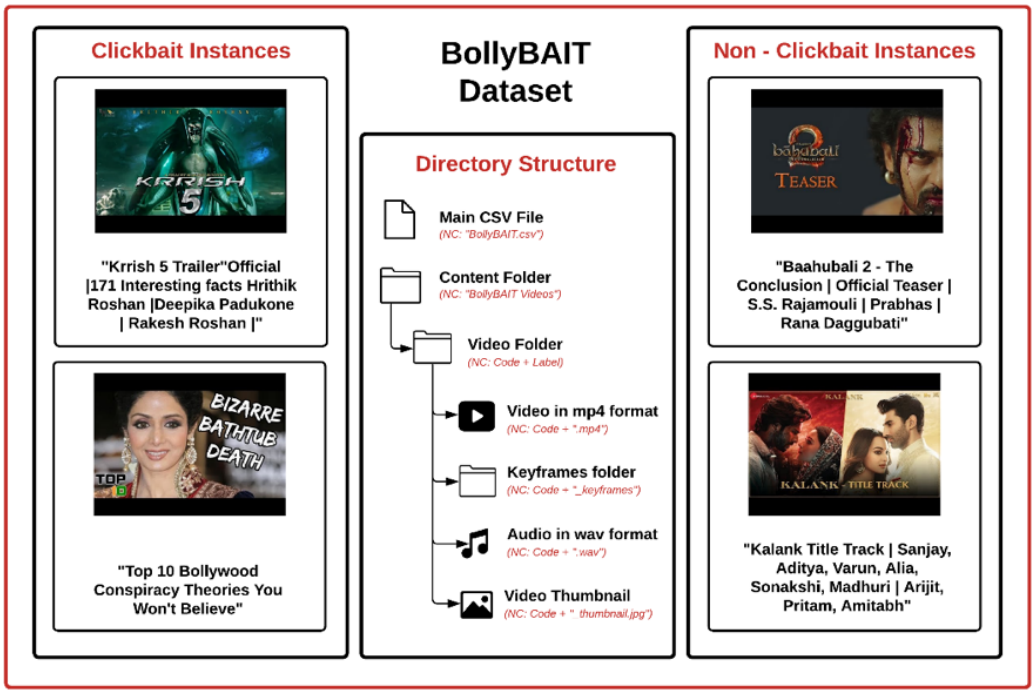


Figure 2: BollyBAIT dataset instances and folder structure

The features of our dataset include: (i) Multilingual: The videos in our dataset belong to many languages such as Hindi, English etc. and therefore provide an opportunity to develop a language-independent model for clickbait detection (ii) Manual Annotation: All the videos have been manually annotated and reviewed by multiple reviewers which allowed us to incorporate categories of clickbait that examine the intricate relationship between the video and its title/thumbnail, currently only possible by human examination such as if the video exploits the curiosity gap of the user and contains false promises or not (iii) Multiple Modalities: In labelling a video as clickbait or not, we have considered four different modalities available: (a) audio of the video, (b) video, (c) text i.e. title of video, and (d) image i.e., thumbnail and extracted features from all of them, often converting one modality into the other (iv) Encompasses a broad definition of clickbait: Our dataset employs a varied and broad definition of clickbait by the virtue of five different categories of clickbaits considered.

4 PROPOSED METHODOLOGY

In this section, we discuss the proposed YouTube clickbait detection framework in detail. The pipeline of the proposed framework for CPDM is composed of the following steps as shown in Figure 3: (i) Pre-processing: This involves extraction of video and associated thumbnail image, title, audio, and keyframes (ii) Feature extraction: Features are extracted from the teaser text (i.e., the Title) and image (i.e., the Thumbnail) of the

video using the pre-trained Bert-Base and Resnet-50 models respectively (iii) Feature set mining: The 6 different feature sets obtained from different modalities are mined using various deep learning techniques (iv) Fusion: Finally, the results of all the deep learning models are combined to predict whether a video is clickbait or not. Each of the steps in the proposed framework are discussed in detail in the following sub-sections.

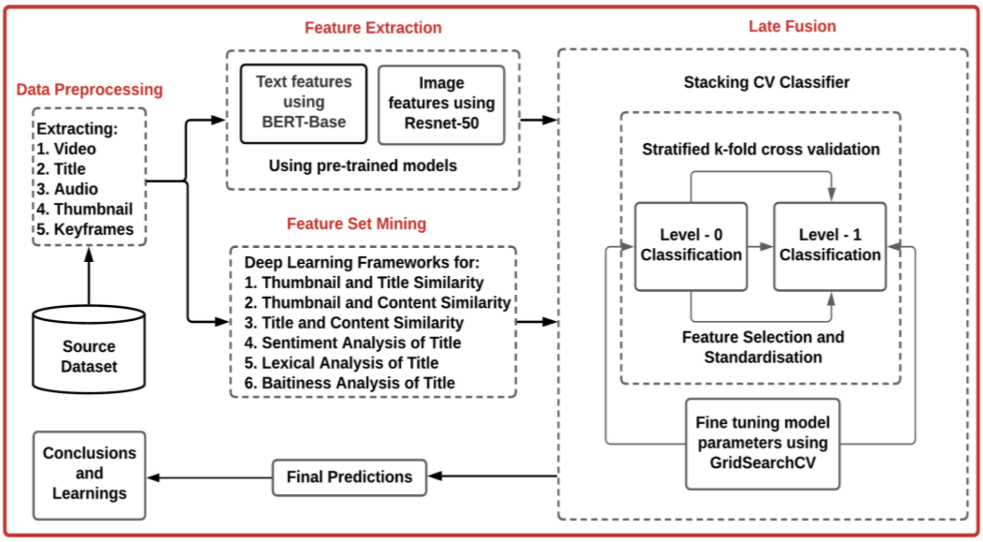


Figure 3: Clickbait Prevention and Detection Framework (CPDM)

ALGORITHM 1: Clickbait Detection and Prevention Framework (CPDM)

for each video in dataset, do
- extract video, audio, thumbnail, title and keyframes
- extract title text features using BERT-Base and thumbnail image features using Resnet-50 models
- perform feature set mining to extract video features, do
  - generate a graph representation for the video to determine title-content disparity
  - generate image caption for thumbnail and compare with title
  - transcribe the audio content and compare with title
  - perform lexical, baitiness and sentiment analysis on the video title
- train and test a stacking classifier on these features with manually annotated label for the video
- return predicted classification label for the video: clickbait or non-clickbait

end

4.1 Data Pre-processing

Video components and descriptors of various modalities i.e., thumbnail, video, audio and title text in the source dataset were scraped using the Python YouTube Data API Library [49]. Subsequently, video keyframes which are the descriptive and essential frames of a video were extracted using the FFMPEG open-source program [50][35] and the video's audio was obtained as a .wav file using the YouTube-dl library [51]. Finally, a

wide variety of features representing the various modalities present in a video were extracted using these various components and fed to the proposed model for training.

4.2 Feature Extraction

Text feature extraction from the teaser text, i.e., the title, was performed using the BERT-Base (Bidirectional Encoder Representations from Transformers-Base) pre-trained model [8]. It is devised to pre-train deep bidirectional representations from an unsupervised text by simultaneously conditioning on both left and right context. Consequently, the BERT Base model can be fine-tuned by adding merely one supplementary output layer to create classifiers for a broad array of NLP assignments such as Sentiment Analysis [14] and Next Sentence Prediction Tasks [33]. In addition, the BERT-Base model can extract conceptualized word embeddings for the text from the title to get the pooled output of 768 units as the feature vector. Image features from the thumbnail associated with the video were extracted using ResNet50 [13] which is a 50-layers deep convolutional neural network trained on more than a million images from the ImageNet database [10], and can classify photos into a thousand target categories. Consequently, the network has acquired deep characteristic representations for an extensive collection of images and, in most cases, can be used without any fine-tuning. Accordingly, it was used to extract image features from the image to get the pooled 2048 units as the feature vector.

4.3 Feature Set Mining

Six prominent feature sets were mined based on thumbnail, title, and video content using various deep learning techniques as follows:

- Content-Thumbnail Disparity: It uses a graph-based deep neural network to produce a vectorized representation of the disparity between the video's thumbnail and the actual content, i.e., the frames of the video.
- Thumbnail-Title Disparity: A text caption for the thumbnail generated using an image captioning neural network was compared with the video's title to get a representation of the disparity between the two texts.
- Audio-Title Disparity: It uses an elaborate model to process audio, transcribe it and then compare the transcribed text to video’s title.
- Sentiment Analysis of Title: The sentiment score of the title was calculated based on its negativity/positivity/neutrality and intensity.
- Lexical Analysis of Title: It is a count-based feature quantifying the presence of lexical cues like capitalized letters or words, excessive punctuations etc. to grab user’s attention.
- Clickbaitiness of Title: It is a count-based feature used to quantify the presence of various enticing and curiosity-generating words and phrases in title which ultimately lure the user to click on the video.

4.3.1 Content-Thumbnail Disparity Module

To analyse the similarity between the video's thumbnail and content, we constructed a framework based on the classification of videos represented as graphs labelled clickbait or non-clickbait. Each video has a weighted directed graph associated with it, which comprises of a root node - the thumbnail, and the children nodes – the keyframes. Each edge from the root to a child captures the resemblance between the two nodes, hence the

graph captures the correlation between the thumbnail and video content. A sample graph has been shown in Figure 4 wherein the thumbnail image is the root node connected to all the child nodes, i.e., the keyframes.

Figure 4: Example of a clickbait video with thumbnail and content disparity

The weight of each edge is calculated as the summation of the following three measures of similarity between the thumbnail and that keyframe: (1) Facial recognition, (2) Object recognition, and (3) Visual image similarity, which have been explained as follows:

1. Facial recognition: It is an empirical expression representing the intersection of faces detected in both keyframe and the thumbnail. Faces present in the thumbnail (F_{TH}) were detected using the DNN Face Detector model in the OpenCV library which is based on the Single-Shot Multibox Detector and the facial embeddings for the same were stored [22]. Then, while processing each of the keyframes, all the faces detected in the keyframe (F_{KF}) using the same MultiBox Detector technique were matched against all those in the thumbnail using the stored facial embeddings for each of the thumbnail, to obtain the number of common faces (CF). This matching is computed by evaluating the Euclidean distance between the two candidate embeddings. Subsequently, the weight was obtained by calculating the relative number of matched faces against the total number of faces present in the keyframe and the thumbnail. The mathematical definition of the weight (W_1) based on this measure is defined in Eq. 1. It is considered to be undefined, when there are no faces found in both the thumbnail and keyframe, since in such a situation, this weight holds no meaning.

$$W_1 = \begin{cases} \frac{2\,CF}{F_{TH} + F_{KF}}, & F_{TH} + F_{KF} > 0 \\ undefined, & otherwise \end{cases} \quad (1)$$

2. Object Recognition: It is an expression representing the convergence of objects present in both the thumbnail and the keyframe. For this purpose, objects were detected using the RetinaNet object detection model [36] and the lists of objects detected in the thumbnail image (O_{TH}) and the keyframes (O_{KF}) were extracted. Then, thumbnail and keyframe lists were matched to get the count of common objects (CO) in the two lists and the subsequent weight was defined as this number against the total number of objects present in the keyframe and the thumbnail. This weight (W_2) has been defined in Eq. 2, and is undefined when no object was present in both the thumbnail and the keyframe.

$$W_2 = \begin{cases} \frac{2\,CO}{O_{TH} + O_{KF}}, & O_{TH} + O_{KF} > 0 \\ undefined, & otherwise \end{cases} \quad (2)$$

3. Visual Image Similarity: This similarity is obtained by comparing the low-level image features that are dependent on the pixels of the image. For this purpose, colour histogram vectors for both the images were constructed by quantifying the colours within the images and enumerating the number of pixels of each colour. Then from these colour histograms, vectors of 3 units – containing the summation, mean and standard deviation of the bins were obtained. The Manhattan distance (d) was used to calculate the distance between these two image vectors which was used as the measure for visual histogram-based similarity. For further elaboration and quantification of this distance, please refer [30]. Finally, we normalise the value of d to obtain this weight (W_3), such that the weight lies in the range of [0,1].

Finally, the mean of these three weights (W_1, W_2 and W_3) is taken to be the edge weight ($\overline{W}$) between the thumbnail and the keyframe, as shown in Eq. 3. Each of these weights lie in the range [0,1] and therefore their mean also lies in the same range. In the situation that either or both weights W_1 or W_2 are undefined, the mean is taken of the remaining defined weights, and the undefined weights are ignored, as they hold no significance.

$$\overline{W} = \frac{W_1 + W_2 + W_3}{3} \quad (3)$$

The video thumbnail and keyframes were processed and corresponding edge weights were calculated and assigned to create a representative graph for each video. Each node in the graph is also assigned its own set of features which is the feature vector extracted using the Resnet-50 pre-trained model of the nodal image. The algorithm used for the supervised classification of these graphs is the Deep Graph Convolutional Neural Network Algorithm [3]. Graph CNN is commonly used in processing of videos as they vary in size and contain complex temporal and spatial interdependencies [20]. The architecture of the graph classification neural network is shown in Figure 5 which is used to extract a pooled output of 16 units from the last layer.

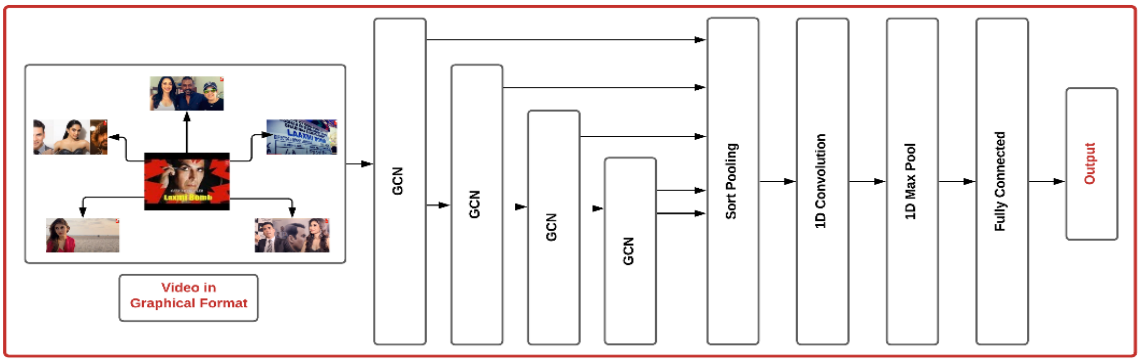


Figure 5: Deep Neural Network based on Graph CNN for identifying thumbnail and content disparity

4.3.2 Title-Thumbnail Disparity Module

Image captioning has been used to capture the textual description of the thumbnail and compare it with the title. Figure 6 shows a clickbait video with the title "new" and a thumbnail depicting a certificate of a movie that has been released. The title doesn't describe what the video contains, neither does it align with the thumbnail itself. This creates a curiosity gap and misleads the viewer into assuming it is a trailer for a new movie.

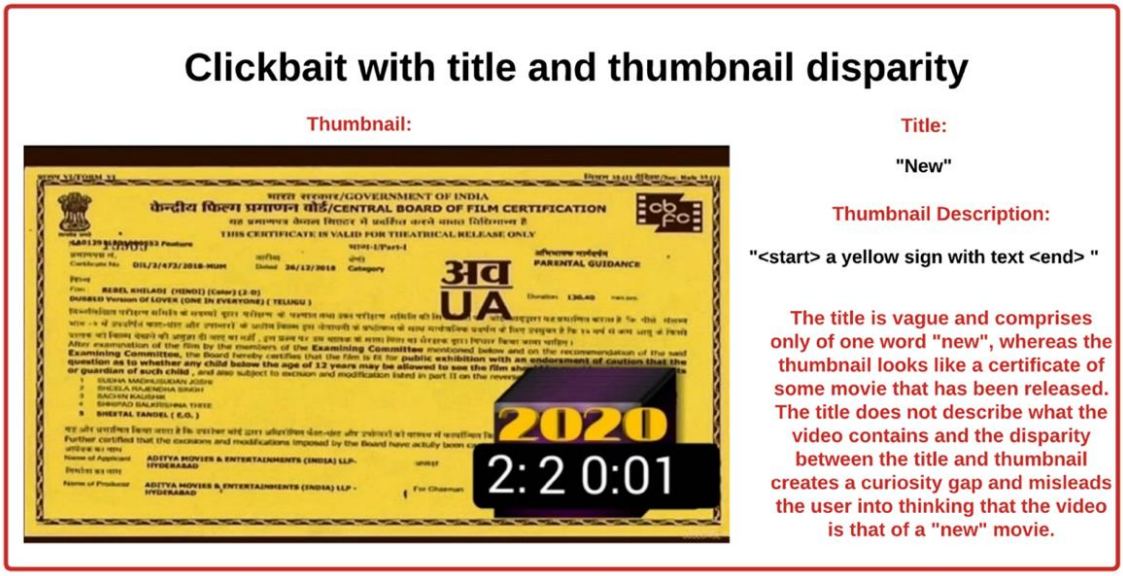


Figure 6: Example of a clickbait video with title and thumbnail disparity

The MS COCO (i.e., Microsoft Common Objects in Context) dataset [19] for image captioning was used to train a neural network-based model for describing for the video thumbnails. It is a large-scale dataset consisting of >300K images with five captions per image. The model contains a CNN encoder which is given images for the task, and its output is fed into the RNN decoder, which gives the generated caption for the image. After

getting the predicted caption for a thumbnail, it was compared with the video title that had been translated to English using the Google Trans Python library [52]. Figure 7 shows the architecture of the model used to caption the images. Afterwards, the comparison between the generated caption and video title was made by employing Cosine Matching [53] on three different measures: (i) on plain text, (ii) on the text pre-processed using NLTK library [4] (iii) on GLOVE [54] embeddings. A feature vector of three units is obtained from this component.

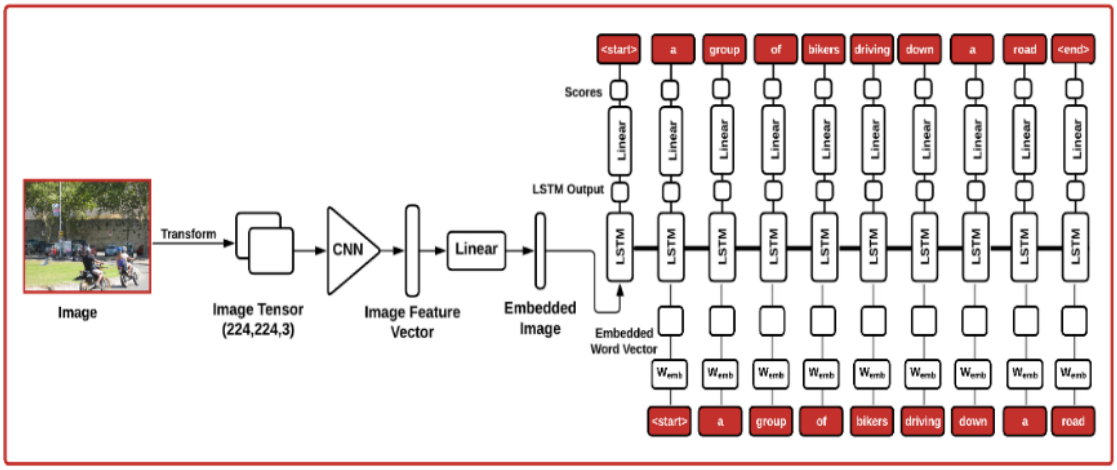


Figure 7: Layers of the deep neural network used for captioning of thumbnail images

4.3.3 Title-Content Disparity Module

This feature component is based on the correlation between the title, i.e., the promised content and the genuine content of the video. Figure 8 shows an example of a clickbait video wherein the title suggests that Ankita Lokhande adopted a baby named Sushant. However, nothing of this description is mentioned in the audio that was transcribed. This transcription is compared with the title to evaluate the similarity between the title and the video content. The audio was converted to text using the Speech to Text Google Speech API [55]. The transcribed text was then translated to English using the Google Trans Python library as the dataset consists of multilingual videos. Consequently, the comparison between the transcribed audio text and the title was made by employing Cosine Matching, as employed in Section 4.3.2.

Figure 8: Example of a clickbait video with title and content similarity

4.3.4 Sentiment Analysis of the Title

A helpful attribute that can give important information about the video is the sentiment of its title [29]. Figure 9 shows examples of clickbait video titles using words indicating strong feelings to entice the user into clicking on them. A strong sense of emotion is often used to draw the viewer to click on a video by exploiting the emotional curiosity gap [39]. We have extracted four sentiment scores for each video: positive valence, negative valence, neutral valence, and compound score. A positive valence score (V_1) is a measure of the positive sentiment in the title; similarly, negative (V_2) and neutral scores (V_3) measure those respective sentiments in the title and each of their values lie in the range [0,1]. These scores are computed using heuristics such as punctuation, capitalization, presence of intensifiers, polarity shifting, and polarity negation. A compound score (C) is a measure of the intensity of sensation in the title which is obtained by summing the valence score of each word and applying normalization so that the value lies in [-1,1]. These scores are quantified by the following equations:

$$\sum_{i=1}^{3} V_i = 1 \tag{4}$$

$$C = \frac{x}{\sqrt{x^2 + \alpha}} \tag{5}$$

where x: sum of valence scores of constituent words, and α: normalization constant

To calculate these scores, we have used the VADER (Valence Aware Dictionary for Sentiment Reasoning) [16], a model used for text sentiment analysis that is sensitive to both polarity (positive/negative) and intensity (strength) of emotion. VADER sentimental analysis relies on a dictionary that maps lexical features to emotion

intensities known as sentiment scores. Using the VADER deep learning model, a feature vector of 4 units was derived.

Figure 9: Example of clickbait videos exploiting intensity of a particular sentiment

4.3.5 Lexical Analysis of the Title

Clickbait video titles use provoking and attention-grabbing cues [26] to entice the user into clicking on them. The most common way to do this is by applying linguistic clues such as overuse of question marks or exclamation marks. Another common occurrence is the use of capital letters to emphasize the enticing words present in the title. Emojis are used, often in excess, to indicate shock or something funny or emotional. Numbers are used for highlighting a good amount of meaty content inside. Figure 10 shows examples of clickbait titles that use these cues to increase the bait value of their video. We extracted the lexical feature vector representing the number of emojis present in the title, presence of numbers in the title, presence of capitalized words in the title, and count of exclamation marks and question marks. These are listed in Table 1.

Table 1: Lexical Features for analysing the title

Feature	Description
If a number present num(x)	Boolean feature which returns true if a number is present in the title and false otherwise.
If a question q(x)	Boolean feature which returns true if the title contains a question mark and false otherwise.
Count of Emojis e(x)	The count of emojis present in the title.

Feature	Description
Relative Frequency of Capital Letters rf(x)	The ratio of capital letters to total number of letters in the title.
Count of Punctuation Marks pun(x)	Returns the count of punctuation marks such as '.','?',',' etc in the title.

Figure 10: Examples of clickbait videos exploiting lexical cues

4.3.6 Baitiness Analysis of the Title

Another striking peculiarity common in clickbait videos is enticing phrases or words in the title to lure the user and take advantage of the curiosity gap to increase clicks and views. These include abbreviations like "OMG" i.e., oh my god to express surprise, "LOL" i.e., laughing out loud to describe humor, "ROFL" i.e., rolling on the floor laughing to express laughter. The use of celebrity names or pornographic words [2] like "nudes" or phrases that exploit the curiosity gap such as "you wouldn't believe", "shocking", etc. is also widespread. Figure 11 shows different clickbait video titles that use these various cues to gain views and mislead users. A list of clickbait phrases that we identified to be prevalent in clickbait video titles has been listed in Table 2. A feature vector containing the count of these enticing phrases in the title of the video was extracted.

Table 2: Features for baitiness analysis the title

Feature	Description
Number of celebrity mentions c(x)	This feature represents the count of the number of celebrities mentioned in the title. Celebrities here are the names present in the celebrities' dictionary in the appendix.
Count of slangs s(x)	This feature represents the number of slang words present in the title, such as “OMG”, or “WTF”.
Count of porn words p(x)	This feature represents the count of porn words present in the title, such as “hot” or “sexy”.
Number of Bollywood-focused enticing phrases b(x)	This feature returns the count of alluring phrases specific to the Bollywood industry, such as “Casting couch”, or “Nepotism”.
Count of generic luring phrases g(x)	This feature returns the number of captivating phrases present in the title, such as “X Reasons Why” or “You wouldn’t believe”.

Figure 11: Examples of clickbait videos with entrancing titles to grab attention

4.4 Late Fusion

Late fusion is a valuable technique employed while combining prediction scores or feature vectors of different modalities to predict the final output by considering the individual outcomes of all the components or models.

After performing feature set mining and combining all the feature vectors, we performed late fusion by employing a stacking classifier framework consisting of six base classifiers and a meta classifier [9]. Stacking is an ensemble learning method that can be applied for achieving classification via a meta-classifier. The base-level or level-0 classifiers are trained on the complete training set, and then the meta-classifier is trained on the predictions of the level-0 classifier-like features. The base-level model often consists of different machine learning techniques, and therefore stacking ensembles are usually wide-ranging and non-homogeneous. Figure 12 shows the framework of the stacking classifier, with six level-0 classifiers, namely, K-Nearest Neighbours, Gaussian Naive Bayes, Extreme Gradient Boosting Multi-Layer Perceptron, Support Vector Machines and Logistic regression, and the level-1 classifier, Random Forest. We split each dataset into training and test sets (8:2) and carried out 10-fold stratified cross-validation for evaluating the model. We adjusted the parameters to use the top 200 best performing attributes of the training data for classification.

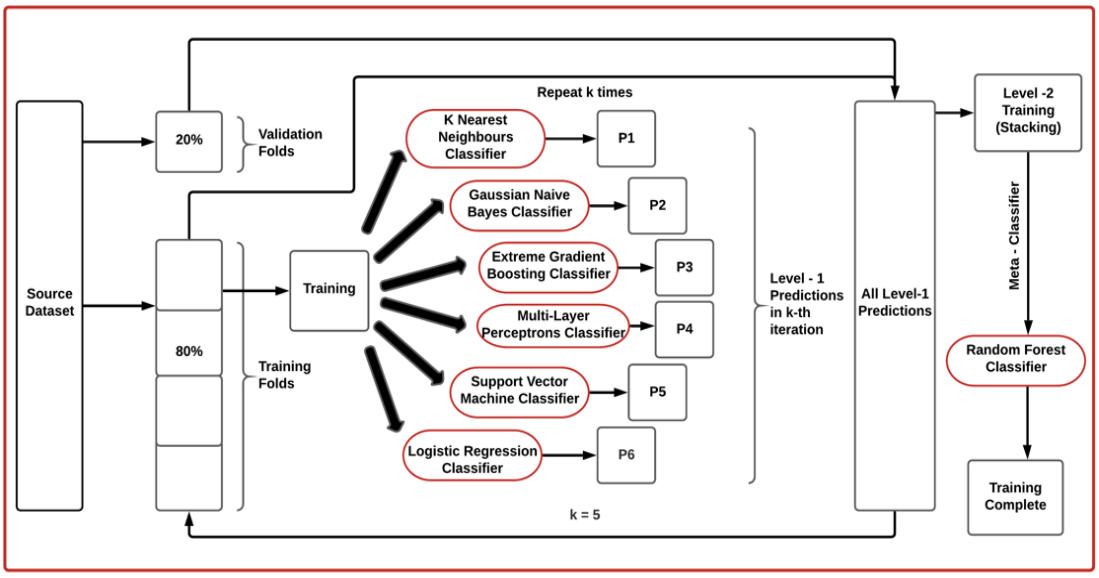


Figure 12: Framework of the stacking classifier

5 EXPERIMENTAL RESULTS

In order to evaluate the performance of the proposed algorithm, we conducted experiments on a Bollywood-focused multilingual YouTube dataset (BollyBAIT) and a general YouTube clickbait-focused dataset, which is the MVD (Misleading Video Dataset). We analysed the performance of the CPDM framework on three aspects, namely classification, feature importance and clickbait categorization on each of these datasets, and these results are showcased below.

5.1 Classification

This section validates our proposed framework by evaluating it on several datasets and reporting the classification results in terms of several evaluation metrics. For experimentation, we curated a Bollywood-

focused multilingual YouTube dataset and combined it with a general YouTube clickbait-focused dataset, which is the MVD (Misleading Video Dataset) [38]. We ran different deep learning algorithms on the two datasets to extract essential features about its telling clickbait characteristics. Subsequently, we trained our stacking classifier on the results from feature mining and evaluated the effects. Table 3 and4 show the fit time (FT), score time (ST), accuracy (ACC), f1-score (F1), precision (PRE), recall (REC), and ROC-AUC (AUC) value for both the BollyBAIT and MVD datasets, respectively, using our stacking classifier and the base models. For both the datasets, the results of the stacking classifier are higher than that of the base models, which proves that the stacking classifier is a better model. All the metrics are the average of the individual metrics over the 10-fold stratified cross-validation run of the classifier. CPDM performed clickbait classification with an accuracy of 92.89% for the BollyBAIT dataset and an accuracy of 95.38% for the MVD dataset. Figure 13 (a) and (b) show the ROC curves for all base models and the Stacking classifier on the BollyBAIT and MVD dataset.

Table 3: Stacking Classifier Results for BollyBAIT

Classifier	FT	ST	ACC	PRE	REC	F1	AUC
Level – 0:							
K-Nearest Neighbours	30.7	0.20	0.930	0.941	0.942	0.941	0.945
Logistic Regression	30.9	0.20	0.924	0.933	0.940	0.936	0.975
Gaussian Naïve Bayes	30.7	0.20	0.920	0.927	0.940	0.932	0.906
Extreme Gradient Boosting	30.7	0.20	0.934	0.950	0.937	0.943	0.967
Multi-Layer Perceptron	30.4	0.20	0.932	0.939	0.947	0.942	0.968
Support Vector Machine	30.5	0.20	0.929	0.941	0.940	0.940	0.976
Level – 1:							
Stacking Classifier	30.4	0.20	0.928	0.941	0.932	0.936	0.976

Table 4: Stacking Classifier Results for MVD

Classifier	FT	ST	ACC	PRE	REC	F1	AUC
Level – 0:							
K-Nearest Neighbours	10.0	10.6	0.885	0.892	0.847	0.868	0.951
Logistic Regression	86.7	00.6	0.943	0.935	0.941	0.937	0.985
Gaussian Naïve Bayes	01.2	01.9	0.791	0.766	0.771	0.766	0.888
Extreme Gradient Boosting	15.7	05.6	0.877	0.890	0.829	0.858	0.956
Multi-Layer Perceptron	18.4	01.1	0.957	0.947	0.959	0.953	0.987
Support Vector Machine	80.4	13.1	0.935	0.938	0.919	0.927	0.982
Level – 1:							
Stacking Classifier	25.1	12.7	0.953	0.947	0.950	0.948	0.983

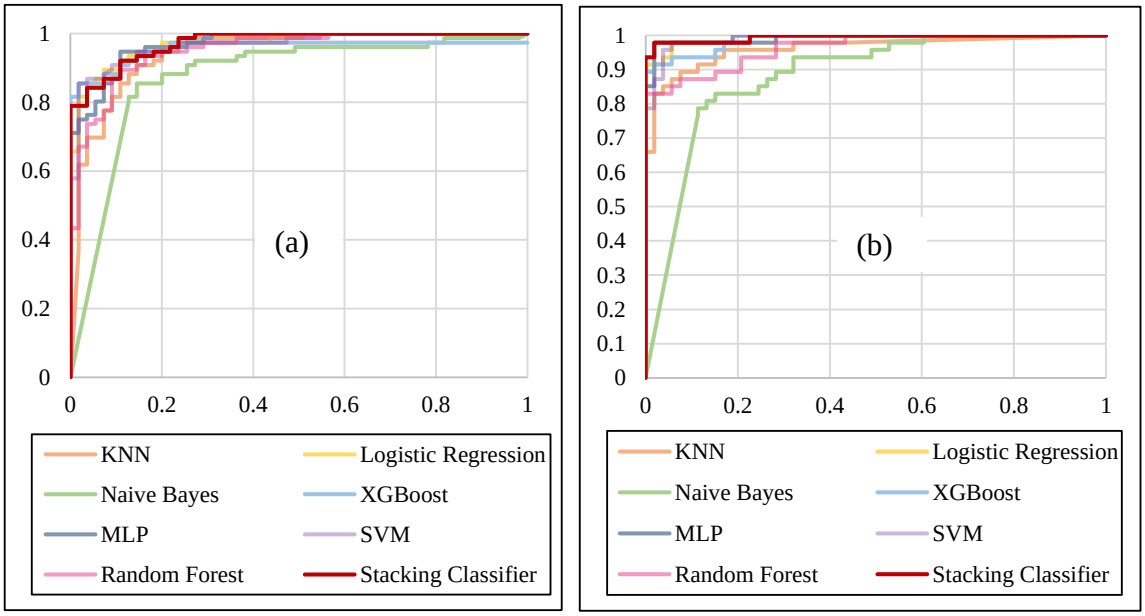


Figure 13: (a) and (b) ROC curve of the BollyBAIT and MVD datasets respectively for the different models used in the stacking classifier.

5.2 Feature Importance

We also conducted experiments to gauge the influence of different feature component combinations on the final classification. Table 5 shows the accuracy, f1-score, precision, recall, and ROC-AUC value using our Stacking Classifier on the self-curated BollyBAIT dataset with individual feature components as well as different feature combinations as shown in the table. All the metrics are the average of the individual metrics over the 10-fold stratified cross-validation run of the classifier.

Table 5: Results for Different Feature Combinations

Feature Group	FT	ST	ACC	PRE	REC	F1	AUC
Individual Feature Sets							
Baitiness Analysis (B)	9.6	0.093	0.732	0.717	0.906	0.799	0.749
Lexical Analysis (L)	12.1	0.094	0.718	0.714	0.872	0.784	0.763
Sentiment Analysis (S)	12.5	0.095	0.653	0.664	0.864	0.734	0.666
BERT Features (BR)	25.5	0.250	0.900	0.898	0.937	0.916	0.964
RESNET Features (RS)	11.9	0.110	0.853	0.867	0.887	0.876	0.913
Title & Thumbnail Disparity Features (TTD)	8.2	0.095	0.640	0.691	0.698	0.693	0.641
Title & Content Disparity Features (TCD)	10.7	0.097	0.658	0.702	0.729	0.714	0.732
Content & Thumbnail Disparity Features (CTD)	12.6	0.097	0.587	0.592	0.950	0.729	0.493
Combination of Feature Sets							
B+L+S	13.2	0.097	0.802	0.790	0.914	0.845	0.849
BR+RS	18.7	0.119	0.899	0.909	0.922	0.915	0.959
TTD+TCD+CTD	14.6	0.100	0.695	0.73	0.757	0.745	0.751

We also evaluated feature importance in terms of the number of features selected for classification, and the results obtained are displayed in Table 6. We found that the highest accuracy of 92.89% was achieved when 825 features were chosen using the ANOVA F-value statistical test, the graph for which is shown in Figure 14.

Table 6: Results for Feature Selection

Selected Features	Accuracy Obtained	Selected Features	Accuracy Obtained
5	79.9934	700	92.0387
50	88.6620	825	92.8964
100	90.0372	900	92.3405
200	90.9650	1000	91.4312
500	90.6643	2000	91.2657

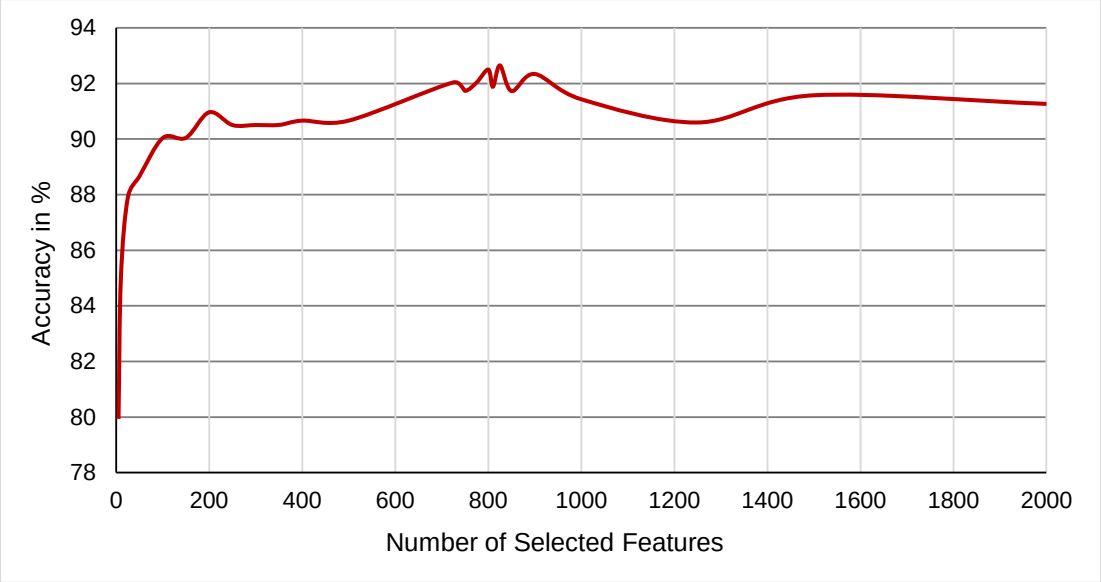


Figure 14: Feature importance graph for the BollyBAIT dataset that peaks at 825 features and 92.89% accuracy

On performing a frequency analysis of all the clickbait titles present in BollyBAIT, we found beguiling words and phrases such as “Official trailer”, “shocking”, “secret”, “Bollywood”, “Shahrukh Khan” that are included in the Lexical analysis of title component, to be prevalent and used frequently, as shown in Figure 15.

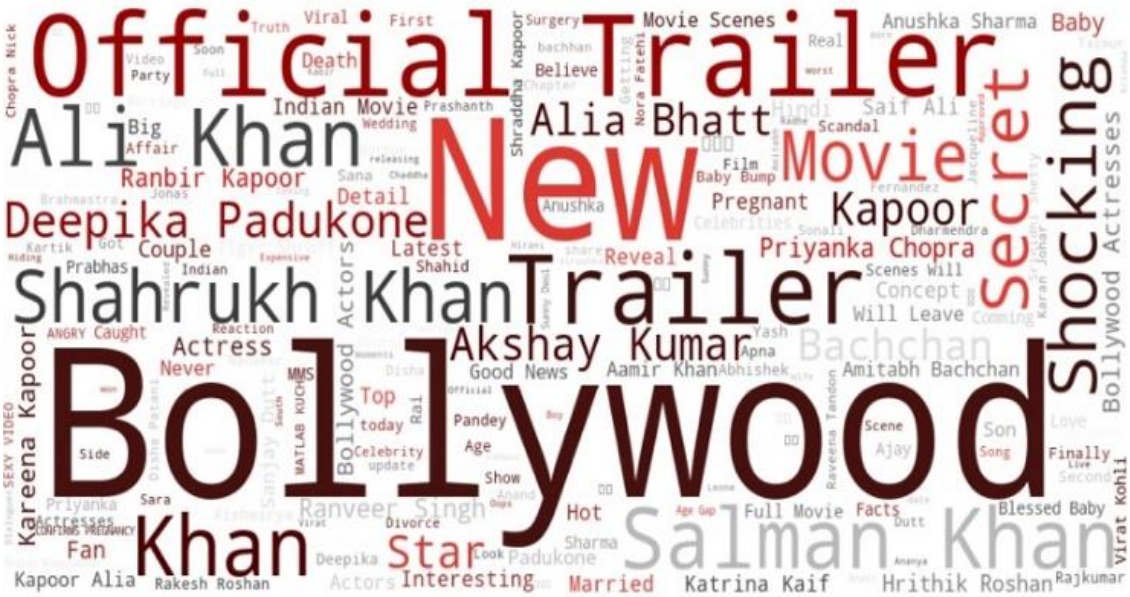


Figure 15: Word cloud of clickbait titles in the BollyBAIT dataset

5.3 Clickbait Categories

The BollyBAIT dataset consists of another set of manual labels describing if the video falls under a particular category of clickbait as defined in Section 1. These different types include: ‘Misleading Video’, ‘Spam Content’, ‘Exaggerated Video’, ‘False Promises’ and ‘Exploits Curiosity Gap’. Clickbait classification for each type of clickbait was done based on these categories. Table 7 shows the accuracy, f1-score, precision, recall, and ROC-AUC value for each clickbait type using our stacking classifier. All the metrics are the average of the individual metrics over the 10-fold stratified cross-validation run of the classifier. Figure 16 shows the correlation matrix for the five different clickbait categories.

Table 7: Results for Different Clickbait Categories

Type of Clickbait	FT	ST	ACC	PRE	REC	F1	AUC
Misleading Video	33.81	0.2041	0.8530	0.723	0.7679	0.7412	0.9193
False Promises	36.50	0.2061	0.8378	0.702	0.7873	0.7384	0.8911
Exaggerated Video	35.65	0.2078	0.8299	0.661	0.6897	0.6706	0.8881
Spam Content	32.74	0.2011	0.8682	0.781	0.8207	0.7976	0.9352
Exploits Curiosity Gap	30.55	0.2028	0.8897	0.716	0.7109	0.7055	0.9193

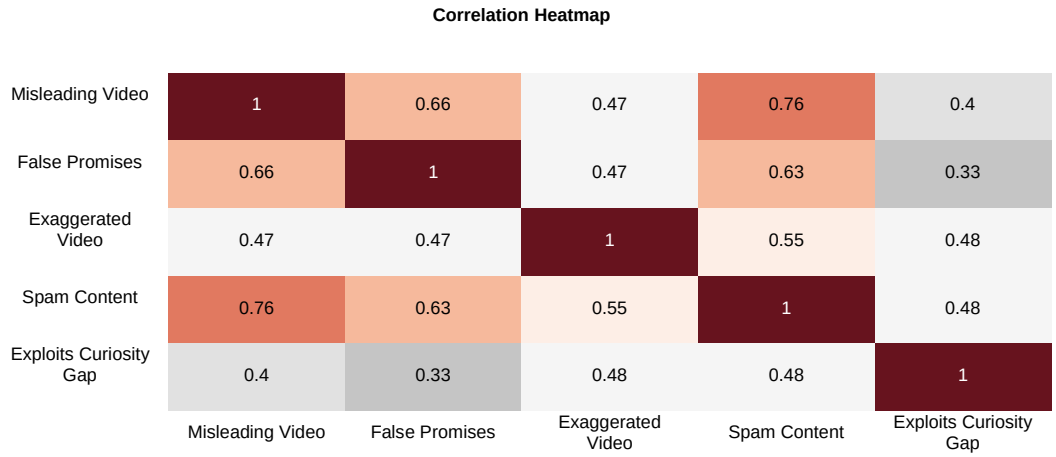


Figure 16: Correlation matrix for the different clickbait categories

6 CONCLUSION

In this paper, we present a novel clickbait prevention and detection model that considers the video's actual audio and video content. We also develop a Bollywood-focused BollyBAIT clickbait dataset containing 1000 videos annotated for five different types of clickbait and a categorical clickbait or non-clickbait label. CPDM leverages video descriptors of other modalities to extract video features effectively and comprehensively for clickbait classification:

- We extract the video frames, audio, title, and other data related to the video.
- We obtain features from the teaser image, i.e., the thumbnail, and the teaser text, i.e., the video's title.
- We perform feature mining across image, text, and audio features by applying deep learning techniques and derive feature sets from different components.
- We implement late fusion on the results of feature set mining for binary classification of YouTube videos as clickbait and non-clickbait using a stacking classifier framework.
- We validate the performance of our CPDM framework by testing it across two clickbait-focused datasets, which gave us an accuracy of 0.953 for the MVD dataset and 0.928 for the self-curated BollyBAIT dataset.

The prime advantage of our work is that because the proposed framework is not dependent on meta-features, it is possible to deploy the model on videos from the moment they are posted. This would decrease the number of clickbait videos circulating by filtering them out from the recommendation feed early, thus improving user experience and engagement. Also, the model considers the actual video content and compares it with the teaser text and image, and the essence of clickbait lies in this comparison. Further work can be done by creating large-scale datasets for clickbait classification and extending CPDM to create a real-time plugin to detect clickbait videos across different video-streaming platforms, not just YouTube.

REFERENCES

[1] Amol Agrawal. 2017. Clickbait detection using deep learning. *Proceedings on 2016 2nd International Conference on Next Generation Computing Technologies, NGCT 2016* October (2017), 268–272. DOI:https://doi.org/10.1109/NGCT.2016.7877426

[2] Cleber S. de Alcântara, Diego Feijó, and Viviane P. Moreira. 2020. Offensive video detection: Dataset and baseline results. *LREC 2020 - 12th International Conference on Language Resources and Evaluation, Conference Proceedings* May (2020), 4309–4319.

[3] Amel Ben Mahjoub and Mohamed Atri. 2019. An efficient end-to-end deep learning architecture for activity classification. *Analog Integrated Circuits and Signal Processing* 99, 1 (2019), 23–32. DOI:https://doi.org/10.1007/s10470-018-1306-2

[4] Steven Bird. 2006. Nltk. January 2006 (2006), 69–72. DOI:https://doi.org/10.3115/1225403.1225421

[5] Prakhar Biyani, Kostas Tsioutsiouliklis, and John Blackmer. 2016. Detecting Clickbaits in News Streams Using Article Informality. *Thirtieth AAAI Conference on Artificial Intelligence* (2016), 94–100.

[6] Abhijnan Chakraborty, Bhargavi Paranjape, Sourya Kakarla, and Niloy Ganguly. 2016. Stop Clickbait: Detecting and preventing clickbaits in online news media. *Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2016* (2016), 9–16. DOI:https://doi.org/10.1109/ASONAM.2016.7752207

[7] Abhijnan Chakraborty, Rajdeep Sarkar, Ayushi Mrigen, and Niloy Ganguly. 2017. Tabloids in the era of social media? Understanding the production and consumption of clickbaits in Twitter. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017). DOI:https://doi.org/10.1145/3134665

[8] Jacob Devlin, Ming Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference* 1, Mlm (2019), 4171–4186.

[9] Saso Džeroski and Bernard Ženko. 2004. Is combining classifiers with stacking better than selecting the best one? *Machine Learning* 54, 3 (2004), 255–273. DOI:https://doi.org/10.1023/B:MACH.0000015881.36452.6e

[10] L. Fei-Fei, J. Deng, and K. Li. 2010. ImageNet: Constructing a large-scale image database. *Journal of Vision* 9, 8 (March 2010), 1037–1037. DOI:https://doi.org/10.1167/9.8.1037

[11] Bilal Ghanem, Paolo Rosso, and Francisco Rangel. 2020. An Emotional Analysis of False Information in Social Media and News Articles. *ACM Transactions on Internet Technology* 20, 2 (2020). DOI:https://doi.org/10.1145/3381750

[12] Yu I. Ha, Jeongmin Kim, Donghyeon Won, Meeyoung Cha, and Jungseock Joo. 2018. Characterizing clickbaits on instagram. *12th International AAAI Conference on Web and Social Media, ICWSM 2018* Icwsm (2018), 92–101.

[13] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 2016-Decem, (2016), 770–778. DOI:https://doi.org/10.1109/CVPR.2016.90

[14] Mickel Hoang, Oskar Alija Bihorac, and Jacobo Rouces. 2019. Aspect-Based Sentiment Analysis using {BERT}. *Proceedings of the 22nd Nordic Conference on Computational Linguistics* (2019), 187–196.

[15] Benjamin D. Horne, Jeppe NØrregaard, and Sibel Adali. 2019. Robust fake news detection over time and attack. *ACM Transactions on Intelligent Systems and Technology* 11, 1 (2019). DOI:https://doi.org/10.1145/3363818

[16] E Hutto, C.J. and Gilbert. 2014. VADER: A Parsimonious Rule-based Model for. *Eighth International AAAI Conference on Weblogs and Social Media* (2014), 18.

[17] Mini Jain, Peya Mowar, Ruchika Goel, and Dinesh K. Vishwakarma. 2021. Clickbait in social media: Detection and analysis of the bait. *2021 55th Annual Conference on Information Sciences and Systems, CISS 2021* (2021). DOI:https://doi.org/10.1109/CISS50987.2021.9400293

[18] Suhaib R. Khater, Oraib H. Al-Sahlee, Daoud M. Daoud, and M. Samir Abou El-Seoud. 2018. Clickbait detection. In *ACM International Conference Proceeding Series*, Association for Computing Machinery, 111–115. DOI:https://doi.org/10.1145/3220267.3220287

[19] Tsung Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. 2014. Microsoft COCO: Common objects in context. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 8693 LNCS, PART 5 (2014), 740–755. DOI:https://doi.org/10.1007/978-3-319-10602-1_48

[20] Hao Liu, Jindong Han, Yanjie Fu, Jingbo Zhou, Xinjiang Lu, and Hui Xiong. 2020. Multi-modal transportation recommendation with unified route representation learning. *Proceedings of the VLDB Endowment* 14, 3 (2020). DOI:https://doi.org/10.14778/3430915.3430924

[21] Qiang Liu, Feng Yu, Shu Wu, and Liang Wang. 2018. Mining significant microblogs for misinformation identification: An attention-based approach. *ACM Transactions on Intelligent Systems and Technology* 9, 5 (2018). DOI:https://doi.org/10.1145/3173458

[22] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng Yang Fu, and Alexander C. Berg. 2016. SSD: Single shot multibox detector. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9905 LNCS, (2016), 21–37. DOI:https://doi.org/10.1007/978-3-319-46448-0_2

[23] Michal Lukasik, Kalina Bontcheva, Trevor Cohn, Arkaitz Zubiaga, Maria Liakata, and Rob Procter. 2019. Gaussian processes for rumour stance classification in social media. *ACM Transactions on Information Systems* 37, 2 (2019). DOI:https://doi.org/10.1145/3295823

[24] Jing Ma, Wei Gao, Shafiq Joty, and Kam Fai Wong. 2020. An Attention-based Rumor Detection Model with Tree-structured Recursive Neural Networks. *ACM Transactions on Intelligent Systems and Technology* 11, 4 (2020). DOI:https://doi.org/10.1145/3391250

[25] Priyanka Meel and Dinesh Kumar Vishwakarma. 2020. Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities. *Expert Systems with Applications 153*. DOI:https://doi.org/10.1016/j.eswa.2019.112986

[26] Bilal Naeem, Aymen Khan, Mirza Omer Beg, and Hasan Mujtaba. 2020. A deep learning framework for clickbait detection on social area network using natural language cues. *Journal of Computational Social Science* 3, 1 (2020), 231–243. DOI:https://doi.org/10.1007/s42001-020-00063-y

[27] Olga Papadopoulou, Markos Zampoglou, Symeon Papadopoulos, and Yiannis Kompatsiaris. 2017. Web video verification using contextual cues. In *MFSec 2017 - Proceedings of the 2nd International Workshop on Multimedia Forensics and Security, co-located with ICMR 2017*, Association for Computing Machinery, Inc, 6–10. DOI:https://doi.org/10.1145/3078897.3080535

[28] Martin Potthast, Tim Gollub, Kristof Komlossy, Sebastian Schuster, Matti Wiegmann, Erika Patricia Garces Fernandez, Matthias Hagen, and Benno Stein. 2018. Crowdsourcing a Large Corpus of Clickbait on Twitter. *Proceedings of the 27th International Conference on Computational Linguistics* (2018), 1498–1507.

[29] Julio Reis, Fabrício Benevenuto, Pedro Olmo, Raquel Prates, Haewoon Kwak, and Jisun An. 2015. Breaking the news: First impressions matter on online news. *Proceedings of the 9th International Conference on Web and Social Media, ICWSM 2015* (2015), 357–366.

[30] Kalyan Roy and Joydeep Mukherjee. 2013. Image Similarity Measure using Color Histogram, Color Coherence Vector, and Sobel Method. *International Journal of Science and Research* 2, 1 (2013), 2319–7064.

[31] Lanyu Shang, Daniel (Yue) Zhang, Michael Wang, Shuyue Lai, and Dong Wang. 2019. Towards reliable online clickbait video detection: A content-agnostic approach. *Knowledge-Based Systems* 182, (2019), 104851. DOI:https://doi.org/10.1016/j.knosys.2019.07.022

[32] Karishma Sharma, Feng Qian, He Jiang, Natali Ruchansky, Ming Zhang, and Yan Liu. 2019. Combating fake news: A survey on identification and mitigation techniques. *ACM Transactions on Intelligent Systems and Technology 10*. DOI:https://doi.org/10.1145/3305260

[33] Wei Shi and Vera Demberg. 2020. Next sentence prediction helps implicit discourse relation classification within and across domains. *EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference* (2020), 5790–5796. DOI:https://doi.org/10.18653/v1/d19-1586

[34] Nguyen Thanh Tam, Matthias Weidlich, Bolong Zheng, Hongzhi Yin, Nguyen Quoc Viet Hung, and Bela Stantic. 2018. From anomaly detection to rumour detection using data streams of social platforms. In *Proceedings of the VLDB Endowment*. DOI:https://doi.org/10.14778/3329772.3329778

[35] Tomar and Suramya. 2006. Converting video formats with FFmpeg. *Linux Journal* 2006, 146 (2006).

[36] Tsung-Yi Lin. 2017. Focal Loss for Dense Object Detection (RetinaNet). *13C-NMR of Natural Products* (2017), 30–33.

[37] Md Main Uddin Rony, Naeemul Hassan, and Mohammad Yousuf. 2017. Diving deep into clickbaits: Who use them to what extents in which topics with what effects? *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2017* July (2017), 232–239. DOI:https://doi.org/10.1145/3110025.3110054

[38] Deepika Varshney and Dinesh Kumar Vishwakarma. 2021. A unified approach for detection of Clickbait videos on YouTube using cognitive evidences. *Applied Intelligence* (2021). DOI:https://doi.org/10.1007/s10489-020-02057-9

[39] Lasya Venneti and Aniket Alam. 2018. How Curiosity can be modeled for a Clickbait Detector. September (2018).

[40] Dong Wang, Tarek Abdelzaher, and Lance Kaplan. 2015. *Social Sensing: Building Reliable Systems on Unreliable Data*. Elsevier Inc. DOI:https://doi.org/10.1016/C2013-0-18808-3

[41] Savvas Zannettou, Sotirios Chatzis, Kostantinos Papadamou, and Michael Sirivianos. 2018. The good, the bad and the bait: Detecting and characterizing clickbait on youtube. *Proceedings - 2018 IEEE Symposium on Security and Privacy Workshops, SPW 2018* March (2018), 63–69. DOI:https://doi.org/10.1109/SPW.2018.00018

[42] Savvas Zannettou, Michael Sirivianos, Jeremy Blackburn, and Nicolas Kourtellis. 2019. The web of false information: Rumors, fake news, hoaxes, clickbait, and various other shenanigans. *Journal of Data and Information Quality* 11, 3 (2019). DOI:https://doi.org/10.1145/3309699

[43] Xinyi Zhou and Reza Zafarani. 2020. A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. *ACM Computing Surveys* 53, 5 (2020). DOI:https://doi.org/10.1145/3395046

[44] Yiwei Zhou. 2017. Clickbait Detection in Tweets Using Self-attentive Network. *arXiv* (2017).

[45] Arkaitz Zubiaga, Ahmet Aker, Kalina Bontcheva, Maria Liakata, and Rob Procter. 2018. Detection and resolution of rumours in social media: A survey. *ACM Computing Surveys 51*. DOI:https://doi.org/10.1145/3161603

[46] Arkaitz Zubiaga and Aiqi Jiang. 2020. Early Detection of Social Media Hoaxes at Scale. *ACM Transactions on the Web* 14, 4 (2020). DOI:https://doi.org/10.1145/3407194

[47] Indian films attract millions globally — and it appears to be growing. Retrieved June 6, 2021 from https://www.cnbc.com/2018/08/03/indian-films-attract-millions-globally-and-it-appears-to-be-growing.html

[48] Bollywood's expanding reach - BBC News. Retrieved June 6, 2021 from https://www.bbc.com/news/world-asia-india-17920845

[49] Python Quickstart | YouTube Data API | Google Developers. Retrieved June 6, 2021 from https://developers.google.com/youtube/v3/quickstart/python

[50] FFmpeg. Retrieved June 6, 2021 from https://www.ffmpeg.org/

[51] youtube_dl · PyPI. Retrieved June 6, 2021 from https://pypi.org/project/youtube_dl/

[52] googletrans · PyPI. Retrieved June 6, 2021 from https://pypi.org/project/googletrans/

[53] Cosine Similarity - an overview | ScienceDirect Topics. Retrieved June 6, 2021 from https://www.sciencedirect.com/topics/computer-science/cosine-similarity

[54] GloVe: Global Vectors for Word Representation. Retrieved June 6, 2021 from https://nlp.stanford.edu/projects/glove/

[55] Speech-to-Text: Automatic Speech Recognition | Google Cloud. Retrieved June 6, 2021 from https://cloud.google.com/speech-to-text

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)
Pune, India. Oct 29-30, 2021

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON) | 978-1-6654-2503-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/SMARTGENCON51891.2021.9645924

Comparative Study of Deep Learning Models for Devanagari OCR

Shilpa Kaur Manocha
Dept. of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
shilpamanocha03@gmail.com

Piyush Tewari
Dept. of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
piyushtewari@dtu.ac.in

Abstract—**Emergence of Deep Learning Approaches have led to major breakthroughs and improvement in image classification tasks. Handwritten character recognition is one such area where the deep learning methods are employed for enhanced performance. This paper reviews the state-of-the-art deep learning methods viz. CNN, ANN, Capsule Network, Residual Network, and Transfer Learning Networks that are pre-trained on the ImageNet dataset— AlexNet, DenseNet, VGG, Inception used for recognition of the handwritten characters of Devanagari script. The paper presents an exhaustive comparison of all mentioned techniques both in terms of qualitative and quantitative analysis. The paper also highlights the challenges, and future directions for Devanagari characters recognition.**

Keywords— Devanagari, Optical Character Recognition (OCR), Deep Learning, Convolutional Neural Network (CNN), Transfer Learning, AlexNet, DenseNet, CapsuleNet, VGG, Inception Net.

I. INTRODUCTION

Optical Character Recognition[1] development has been gaining popularity in recent years, for Devanagari script along with other Indic Scripts. The script serves as a base for over a 100 languages around the world including few popularly used scripts :- Hindi, Marathi, Sanskrit etc. Development of robust OCR systems for Devanagari script will allow us to preserve old manuscripts by converting the physical files to digital formats. This will make the process of storage, retrieval and transfer very convenient.

For recognition of isolated characters from physical documents and images, the first step is image pre-processing that includes steps like binarization, skew removal, slant removal, thresholding and noise removal. Next step is feature extraction i.e. instead of feeding all raw pixels as input to the classifier for classification, several methods are employed to detect important features from the images. For text, these features may be horizontal lines, vertical lines, diagonals, loops, dots etc. Selection of features that can best describe the dataset and their extraction is a crucial step and affects the final recognition accuracy. The classifiers like SVM, KNN take the combination of these features as inputs and classifies them into different classes.

From the early 2000's, the emergence of deep learning [2] approaches have revolutionised the way we perform handwriting recognition and other classification tasks in the field of computer vision & image processing tasks. The use of Deep Models have eliminated the need of explicit feature extraction methods. They are capable of selecting appropriate features depending upon the raw data fed and can also perform the classification tasks. Some of the latest deep learning models employed for Devanagari characters recognition are ANN, CNN, AlexNet, ResNet, VGG, Inception Net, CapsuleNet, DenseNet.

In this paper, a review of the state of the art deep learning approaches for recognition of handwritten Devanagari characters has been presented. The results are made available in the form of tables for the easy comprehension.

The paper is structured into sections as follows:- Section I gives Introduction. Section II highlights the challenges in Devanagari OCR systems, Section III provides explanation of the state-of-the-art Deep Neural Networks and Transfer Learning Models selected for study, Section IV is dedicated to the Comparative Analysis of the mentioned methods on Devanagari dataset on DHCD[20] dataset. Section V gives directions for future work and Section VI gives Conclusion.

II. CHALLENGES IN DEVANAGARI OCR

The Devanagari script [3] has various characters which have similar structures and differ only in minor details like dots, horizontal line, loops etc. This makes the classification of Devanagari characters very challenging. And on top of that different individuals may write the same characters differently due to cursive nature of the script. Table 1. shows few similar characters in Devanagari. Non-Availability of benchmark datasets of Devanagari script is yet another major issue. At present we have very limited number of Devanagari datasets as compared to English or other non-Indic scripts. There is a requirement of large and inclusive Datasets like we have for English. The number of datasets available are very less and small in size for both isolated characters and words, when compared with those available for English.

Fig. 1. Devanagari Script characters

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:58:36 UTC from IEEE Xplore. Restrictions apply.

TABLE I. SIMILAR DEVANAGARI HANDWRITTEN CHARACTERS

daa	kna	Characters differing only in single dot
da	dhaa	Difference only in closed loop
dha	chha	Difference in style of 'Shirorekha'- header line and closed loop
dha	gha	Difference only in loop
bha	tha	Characters differ only in loop and curve

III. DEEP LEARNING ARCHITECTURES FOR DEVANAGARI OCR SYSTEMS

A. ANN/DNN

Artificial Neural Networks [4] are merely a collection of nodes or artificial neurons that resemble the biological neurons in the human brain. These neurons are associated with one other and each connection has a weight associated with it. ANN consists of three kinds of layers: the input layer, intermediate (hidden) layer/s, and the output layer. The weights of each connection are updated by the method of back propagation to improve the accuracy and reduce loss of the classification task. Depending on the number of hidden layers in the model, it can be named as ANN (1 hidden layer) or Deep Neural Network-DNN (multiple hidden layers).

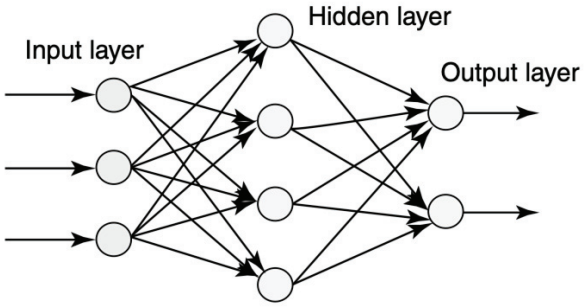


Fig. 2. Artificial Neural Network [4]

B. CNN

Convolutional Neural Networks [5] are popularly used for image classification & object detection tasks. The Network is inspired by the learning model of neurons of human brain and requires very less pre-processing of input images. They consist of three types of layers: - (i) Convolution layers - which perform the operation of convolution of input images with filter or kernel and generates feature maps. (ii) Pooling Layer: - down samples the size of feature maps. It reduces the dimensions of features by performing pooling or aggregation operation. The most commonly and widely used pooling methods are average and max pooling. The convolution layers and pooling layers are together called as Feature Extractors. (iii) Fully Connected layers: are the layers responsible for performing the classification of dataset images into respective classes. They take the features from previous layers as inputs and based on the combination of features learned, it can identify the samples correctly. The model learning happens because of loss function and optimizer. The loss function tells the model its task. e.g., to reduce the mean absolute error between predicted class and actual class label. The optimizer tells the network on how to achieve that aim. Most popular optimizers are Adam and RMSProp that do an excellent job of updating weights in an adaptive manner and thus improving accuracy of model.

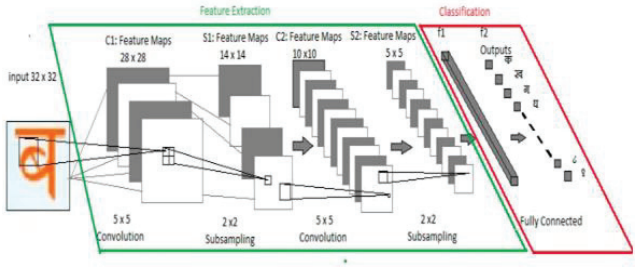


Fig. 3. Convolutional Neural Network [20]

C. AlexNet

This model was developed by *Alex Krizhevsky [8]* for classification of the large sized dataset - ImageNet. The dataset consists around 15million images labelled with 22k classes. The architecture consists of five convolution layers and three fully connected layers for classification of the labelled dataset. The convolutional layers make use of large receptive fields using 11x11 filter size with stride 4. The model incorporates some features over traditional CNN like (i) ReLU non-linearity, (ii) multiple GPUs and (iii) Overlapping pooling for the performance enhancement. Use of ReLU non-linearity makes the architecture faster as compared to use of tanh and sigmoid function. Usage of multiple GPUs makes the training of bigger dataset easier while cutting down the training time. Introduction of overlapping pooling causes a reduction in error and also reduces the chances of overfitting. AlexNet has around 60 million parameters which poses the problem of overfitting. The issue has been tackled using two approaches namely - 'Data Augmentation' and 'Dropout'. Data Augmentation has increased the size of training set by a factor of 1024 by making transformations on the images and provides more varied data. Dropout helps prevent overfitting but also increased the training time needed for convergence of model. The effectiveness of the architecture has been established when it won the ILSVRC Challenge in 2012. It achieved a top-5 error rate of 15.3%. AlexNet outperformed all the other models.

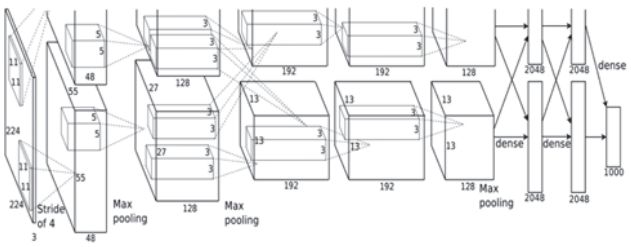


Fig. 4. Architecture of AlexNet [8]

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:58:36 UTC from IEEE Xplore. Restrictions apply.

D. VGG

Unlike AlexNet, VGG [10] model employs use of small size convolutions with filter size of 3x3. The convolution layers make use of small receptive fields(3x3) and stride is fixed to be 1 so as to preserve the spatial resolution. All hidden layers make use of ReLU function. And there are 3 fully connected layers. VGG's performance is better than that of AlexNet in terms of training time and number of parameters. The small convolutions lead to a large number of weight layers and thus enhanced performance but increased computational cost. VGG models support up to 19 layers and are popularly used in image classification & computer vision tasks.

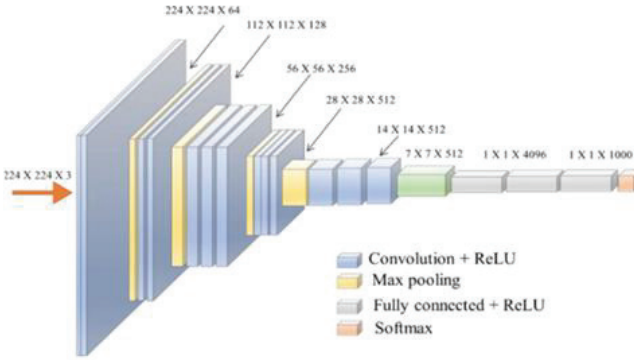


Fig. 5. Architecture of VGG-16[17]

E. Inception Network

It is an extension and improvement of the VGG model and implements multiple filters of sizes of 3x3, 3x1, 1x3, 1x1. The motivation behind the network was to make it deeper [11]. Use of these different filter sizes results in fewer number of parameters and thus reduction in computations as compared to VGG architecture. Inception Blocks are used that consist of different paths to extract information from inputs using different filters. The final output of an inception block consists of combination of all the spatial information features extracted through the different paths. Inception Blocks have fewer parameters and less computation complexity than single 3x3 and 5x5 convolutional layers. Inception Model has several variants- Inception V1, Inception V3, Inception V5.

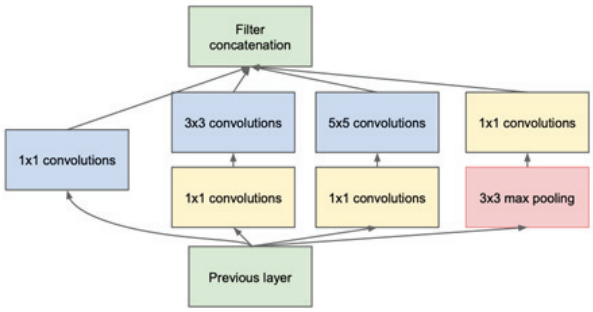


Fig. 6. An Inception Module [11]

F. Residual Newtork (ResNet)

Deep Learning Models have large number of training layers which leads to problems like overfitting and vanishing gradients. When there is an excess number of layers, the model learns too many specific features and is able to solve complex classification problems but may fail to classify simple tasks. The accuracy, thus, starts to saturate and finally degrades. To resolve this problem in deep architectures, Residual Network was developed [6]. The model makes use of Residual function and skip connections to achieve this. The skip connections skip training over few layers and connect input layers directly with the output. This allows the model to memorise Residual Mapping instead of initial mapping and this leads to better performance due to regularization. Figure 5. Shows a residual block. Consider a NN having input as x that approximates the function 'H(x)'. 'R(x)' is the residual function that represents the difference between these two i.e., 'R(x)' = (H(x) - x). Layers in the residual function try to learn the residual function - 'R(x)'. The original function thus becomes F(x) + x. The Residual Network (ResNet) consists of multiple residual blocks.

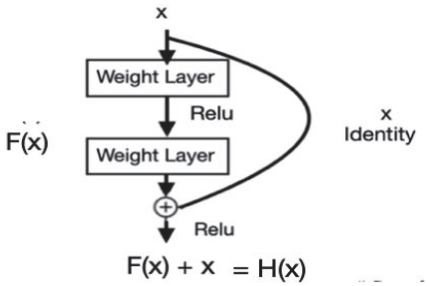


Fig. 7. Residual Block

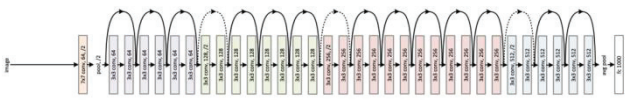


Fig. 8. Residual Network (ResNet-34) [6]

G. DenseNet

DenseNet [9] makes use of feature reuse instead of using deep and wider architectures. The model's layers are very narrow having small number of filters. DenseNets concatenate the input and output feature maps in each layer, instead of summing them up as in other CNN networks. Each layer has access to previous feature maps. Thus, all layers share the collective knowledge about features. To solve the problem of training in deep networks where information path is large from input to output, DenseNets solve this problem by providing each layer the direct access to the gradients obtained from the loss function & the original input image. Since the feature maps sizes added at each layer vary in size, the model makes use of Dense Blocks for concatenation. Each dense block has feature maps having same dimension. The number of filters may vary though. In between these dense blocks are transition layer that performs tasks like batch normalization, convolution and pooling.

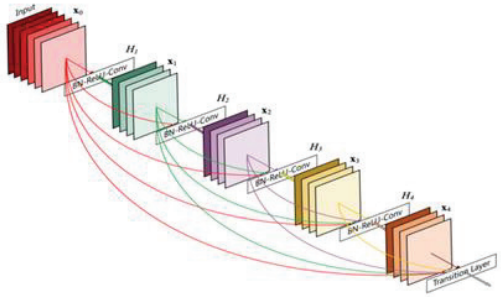


Fig. 9. A 5-layer Dense Block Architecture [9]

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:58:36 UTC from IEEE Xplore. Restrictions apply.

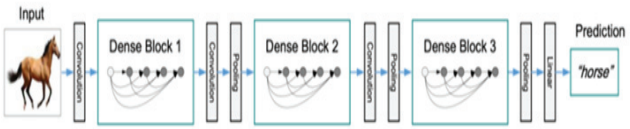


Fig. 10. DenseNet Architecture with 3 dense blocks [9]

H. Capsule Networks

ConvNets make use of max pooling layers. Max Pooling leads to loss of spatial information like translation, pose, orientation etc. These limitations are overcome by Capsule Networks [7]. Capsule Networks are made up of layers of capsules. Capsules accept input vectors instead of scalar as with the case of other Neural networks. Input vectors consist of probabilities in the form of length vectors and the directions of the features that are detected. Lower capsules pass on the information about features, their directions and probabilities to the next higher-level capsules which resemble the overall features of higher layer capsules. This method of updating scalar weights of input vectors is known as dynamic routing by agreement, as opposed to back propagation that is used in CNN. Thus, they can retain spatial information of features as well. The Network consists of two parts: 1) Encoder and 2) Decoder. The encoder consists of 3 types of layers. (i) Input Convolution layer that detects basic features. (ii) Capsule Layers: contain outputs from first layer and produce combination of features and route the features to next appropriate capsule layer by method of dynamic routing by agreement. (iii) Character layer: has number of capsules equal to the number of classes to be identified by the network. Outputs from this layer are fed to Decoder. The decoder consists of fully connected layer responsible for recognition of classes. Decoder also performs reconstruction of images.

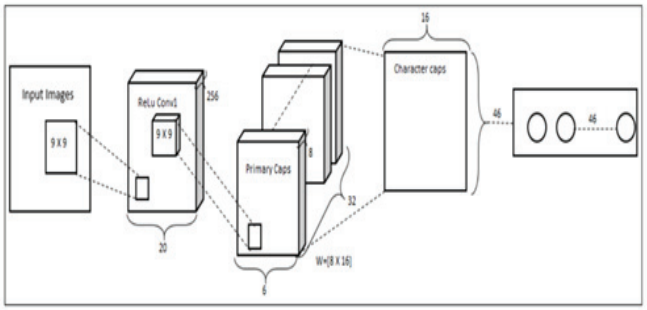


Fig. 11. Architecture of Encoder Part of Capsule Network [15]

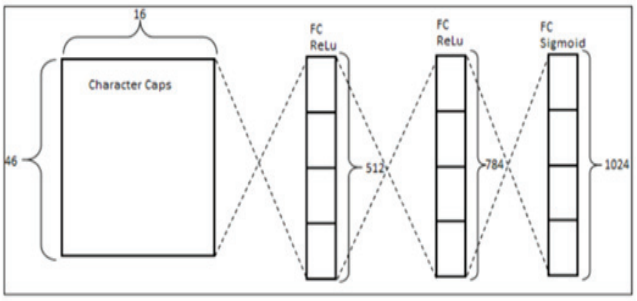


Fig. 12. Architecture of Decoder Part of Capsule Network [15]

IV. Comparative Analysis

A. Dataset

For Analysis of the above-mentioned approaches, results on the DHCD dataset have been considered. DHCD dataset was proposed by Acharya et. al [20] and is used for analysis and comparison. The dataset comprises of total 46-character classes - 36 Devanagari characters and 10 numerals. Each class has 2000 images of data. i.e., the dataset is balanced with a total of 92000 images.

B. Evaluation Parameters

For recognition of handwritten characters, the most widely used evaluation parameters are accuracy and time required for training the model.

1) Classification Accuracy: It is the ratio of correct predictions to the total number of input samples. It should be as high as possible.

2) Number of Epochs: An epoch means training the NN with all training data. For training a network, generally, multiple number of epochs are required so that the model can learn, and extract features needed to classify the classes with good classification accuracy. Lesser number of epochs required for training the model implies that it learns faster and thus, is the desired outcome.

For Quantitative Analysis, models that are trained from scratch and the transfer learning models both are considered. Transfer Learning is a technique that allows one to perform a classification task by use of a model that can has been trained to perform classification on a different but related task. Models have been pre-trained on ImageNet dataset for classification of about 1000 classes are used. These pre trained models are used with their weights and features learned as fixed feature extractors and are fed to fully connected layers of CNN for classification of Devanagari characters.

For Qualitative Analysis, different deep learning techniques discussed in this paper are represented in table 3, for better comprehension and readability. The table mentions key issues addressed and limitations of each model for recognition of Devanagari characters.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:58:36 UTC from IEEE Xplore. Restrictions apply.

TABLE II. RECOGNITION RESULTS OF DEEP LEARNING APPROACHES FOR DEVANAGARI OCR.

Model	Training / Feature Extraction Method	Classifier Architecture	Classification Accuracy	Number of Epochs	References
Model-1 ANN	Models trained from scratch	Legendre moment based feature extraction & classification using ANN	98.25%	-	[16]
Model-2 CNN		CNN (4 convolutional layers)	98.47% and 98.26% on deep and shallow models	50	[20]
Model-3 ResNet		ResNet-34	98.73%	5	[18]
		ResNet-50	99.35%	5	[18]
Model 4- Capsule Net		Capsule Net	99.075	50	[15]
Model-5 AlexNet	Pre-Trained Models	AlexNet (8 layers)	98%	3	[17]
Model-6 DenseNet		DenseNet-121	89%	7	[17]
		DenseNet-201	90%	6	[17]
Model-7 VGG		VGG-11	99%	8	[17]
		VGG-16	98%	3	[17]
		VGG-19	98%	3	[17]
Model-8 Inception		Inception- V3	99%	1	[17]

TABLE III. QUALITATIVE ANALYSIS OF DEEP LEARNING APPROACHES FOR DEVANAGARI HCR.

Neural Networks					
Model	**Year Proposed**	**Major Contributions**	**Key Issues Addressed**	**Limitations**	**References**
ANN	1957	Only 1 hidden layer	simple to build network	requires explicit feature extractor Performs only classification Large number of parameters	[4],[16]
CNN	1980	Automatic extraction of appropriate features. Highly customizable - number of layers, number of filters and kernel size	Significant improvement in training time & accuracy as compared to traditional classification methods.	Large dataset requirement for training Overfitting due to smaller dataset or deep architecture Doesn't classify well on test image samples having translation, pose, rotation ,scaling etc. Loss of spatial information because of pooling layers.	[5],[20]
AlexNet	2012	First Model developed for attaining high accuracies on very large datasets by making use of multiple GPUs , overlapping pooling and use of ReLU non-linearity.	Features learned from training on ImageNet dataset can be used for classification of small datasets of Devanagari directly.	dropping of any one layer from the network architecture impacts the model performance immensely. Requirement of multiple GPUs .	[8],[17]

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:58:36 UTC from IEEE Xplore. Restrictions apply.

VGG (VGG11, VGG16, VGG19)	2014	uses small sized convolution filters and stride of 1 to retain spatial features.	huge enhancement in performance and training time over AlexNet	increased computational cost due to small convolution filters and large number of weight layers.	[10],[17]
Inception Net (GoogLeNet)	2015	use of multiple sized filters in an inception module which leads to fewer parameters than VGG model.	reduced computational cost as compared to VGG networks.	deep network Visually complex network	[11],[17]
ResNet (ResNet-34, ResNet-50)	2015	Makes use of Residual functions or skip/short connections to avoid overfitting.	Solves problem of vanishing gradient and overfitting on increase in number of layers (unlike in CNN)	training the model from scratch takes time due to large number of layers involved.	[6],[18]
DenseNet (DenseNet -121, DenseNet-201)	2017	Simplifies the connectivity patterns between layers by making use of Dense Blocks	Requires fewer parameters as compared to traditional CNN Prediction time is low when pertained model is used.	Visualisation becomes difficult because of dense number of connections.	[9],[17]
Capsule Net	2017	Network made up of Capsules instead of neurons Capsules aggregate relevant features from one object in the image. Separate encoder and decoder parts.	Invariant to translation, rotation in the dataset through use of equivariant mapping Provides better results for overlapping characters in an image Decoder at the end of network allows to regenerate images with reduced noise & more clarity	very high training time and computational cost. Training of a single epoch for a medium sized dataset takes several hours. The model is simple, but the time required for computations is large owing to the inner looping by the routing algorithm.	[7],[19]

V. Future Directions

1. Development of a benchmark handwritten characters dataset containing all primary characters, compound characters and numerals in handwritten form. The dataset must incorporate handwriting of different individuals from a large community, so that the dataset can have more variance in terms of translation, rotation, illumination and it must consist of a very large sample size.
2. Propose a novel feature extraction method for selection of appropriate features of Devanagari characters that can allow classifier to distinguish similar characters and improve accuracy for the cursive Devanagari characters.
3. Exploration of Xception Net, XGB Classifier for recognition of Devanagari characters and exploration of XGBoost Algorithm for improvement of classification of handwritten Devanagari characters.

VI. Conclusion

This paper represents an exhaustive comparison of state-of-the-art deep learning approaches used for classification of Devanagari handwritten characters. Models selected for study include those that have been trained from scratch and those Pre-trained on the ImageNet Dataset. The results models have been compared based on the recognition accuracy and training time required. Qualitative analysis has also been done for all models and the key features, merits and limitations have been discussed in a tabular form. Based on the survey of the methods adopted till now and challenges faced, future directions have also been pointed out. Thus, the paper provides an insight into various Devanagari handwritten character recognition methods and the readers will benefit from the presented analysis for the selection of appropriate model/s as per their requirement.

References

[1] Chaudhuri A., Mandaviya K., Badelia P., Ghosh S.K. (2017) Optical Character Recognition Systems. In: Optical Character Recognition Systems for Different Languages with Soft Computing. Studies in Fuzziness and Soft Computing, vol 352. Springer, Cham.

[2] LeCun, Y., Bengio, Y. & Hinton, G. "Deep learning", Nature 521, 436–444 , 2015.

[3] Bright, W.,. "The devanagari script.", The world's writing systems, pp.384-390,1996.

[4] Abraham, Ajith. "Artificial neural networks." *Handbook of measuring system design* (2005).

[5] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," 2017 International Conference on Engineering and Technology (ICET), pp. 1-6, 2017.

[6] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In "*Proceedings of the IEEE conference on computer vision and pattern recognition* "(pp. 770-778)**.**

[7] Sabour, S., Frosst, N., & Hinton, G. E. (2017). Dynamic routing between capsules. *arXiv preprint arXiv:1710.09829***.**

[8] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84-90.

[9] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In "*Proceedings of the IEEE conference on computer vision and pattern recognition"* (pp. 4700-4708).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:58:36 UTC from IEEE Xplore. Restrictions apply.

[10] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

[11] Szegedy, Christian, et al. "Going deeper with convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

[12] P. K. Sonawane and S. Shelke, "Handwritten Devanagari Character Classification using Deep Learning.," 2018 International Conference on Information , Communication, Engineering and Technology (ICICET), pp. 1-4, 2018.

[13] R. Kumar, A. Kumar, and P. Ahmed, A Benchmark Dataset for Devanagari Document Recognition Research, WSEAS Press, Lemesos, Cyprus, 2013.

[14] A. Mohite and S. Shelke, "Handwritten Devanagari Character Recognition using Convolutional Neural Network," 2018 4th International Conference for Convergence in Technology (I2CT), pp. 1-4, 2018.

[15] S. Gupta and R. K. Mohapatra, "Performance Improvement in Handwritten Devanagari Character Classification," 2019 Women Institute of Technology Conference on Electrical and Computer Engineering (WITCON ECE), 2019, pp. 60-64.

[16] K. V. Kale, S. V. Chavan, M. M. Kazi and Y. S. Rode, "Handwritten Devanagari Compound Character Recognition Using Legendre Moment: An Artificial Neural Network Approach," 2013 International Symposium on Computational and Business Intelligence, 2013, pp. 274-278.

[17] Aneja, Nagender, and Sandhya Aneja. "Transfer learning using CNN for handwritten devanagari character recognition." *2019 1st International Conference on Advances in Information Technology (ICAIT)*. IEEE, 2019.

[18] Mhapsekar M., Mhapsekar P., Mhatre A., Sawant V. (2020) Implementation of Residual Network (ResNet) for Devanagari Handwritten Character Recognition. In: Vasudevan H., Michalas A., Shekokar N., Narvekar M. (eds) Advanced Computing Technologies and Applications. Algorithms for Intelligent Systems. Springer, Singapore.

[19] Sukhadiya J., Suba Y., D'silva M. (2020) Devanagari Character Classification Using Capsule Network. In: Abraham A., Cherukuri A., Melin P., Gandhi N. (eds) Intelligent Systems Design and Applications. ISDA 2018. Advances in Intelligent Systems and Computing, vol 940. Springer, Cham.

[20] S. Acharya, A. K. Pant and P. K. Gyawali, "Deep learning based large scale handwritten Devanagari character recognition," 2015 9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA), pp. 1-6,2015.

[21] Deore, S.P., Pravin, A. Devanagari Handwritten Character Recognition using fine-tuned Deep Convolutional Neural Network on trivial dataset. *Sādhanā* **45,** 243 (2020).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:58:36 UTC from IEEE Xplore. Restrictions apply.

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)
Pune, India. Oct 29-30, 2021

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON) | 978-1-6654-2503-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/SMARTGENCON51891.2021.9645784

Comparative Study Of Impact Of COVID-19 Lockdown On AQI Parameters Across Urban India

Himanshu Nandanwar
2k20/ISY/10(information technology)
Delhi Technological University
Delhi, India
himanshunandanwar9cm0@gmail.com

Anamika Chauhan
Information Technology
Delhi Technological University
Delhi, India
anamika@dce.ac.in

***Abstract*—In the midst of the COVID-19 outbreak, India imposed a state-wise lockdown due to which pollution levels in major cities across the nation decrease drastically. The current study used a variety of pollution metrics (PM10, PM2.5, SO2, NO2, CO, O3, and NH3) to reveal geographical air quality patterns before and during the lockdown. The findings revealed that during a lock-down, air quality increases significantly. India suffered significant economic losses as a result of the shutdown, but it did relieve the environment. The goal of this study is to determine the effects of lockdown on air quality in 26 major Indian cities. A significant reduction in the concentration of PM2.5/PM10 levels is observed as compared to the other pollutant levels. O3 concentration is higher between 2017-2019 as compared to the lockdown phase. Delhi observed the drastic change in pollution level going from the very poor category to the good category. In general, the study is believed to be a useful input to regulatory agencies because it demonstrates that reducing pollution sources can improve air quality. Temporary source control can help cure the environment for a limited time.**

***Index Terms*—AQI, Covid-19, Air pollution, LocKdown, Human-health, Air quality, India, Pollutants.**

I. Introduction

The Indian government, led by Prime Minister Narendra Modi, imposed a 21-day statewide lockdown on March 23, 2020, as a preventive step against the COVID-19 sickness in India, prohibiting mobility for the whole Indian population of 1.38 billion, or 138 crores. And so our lives as Indians would begin in the midst of the epidemic1. People began to notice that the environment, particularly the air quality, had improved considerably almost immediately after the curfew was implemented. So much so that the picturesque Dauladhar Mountains in Himachal Pradesh may be seen from its bordering states.

Coronaviruses are a wide range of viruses that both animals and humans can be infected with. Many coronavirus diseases in humans have been connected from the ordinary cold to more serious disorders such as Middle East Respiratory Syndrome and Severe Acute Respiratory Syndrome (SARS) with respiratory infections2. The newly found coronavirus is the cause of coronavirus disease. COVID-19 was recognized as a public health emergency by the World Health Organization (WHO)[9]. figure 1 shows the covid-19 timeline of world.

The number of cases worldwide is expanding daily. This dataset covers statistics from India and union territories on a universal scale. Air is a necessary component of life. The better our health and well-being, the cleaner the air[3]. However, artificial sources, in particular, pose a considerable risk to air quality.

Air pollution is the release of chemicals into the atmosphere that are harmful to human and world health[4]. It has the potential to be one of humanity's most dangerous threats ever. All of them are affected animals, agriculture, forest, and water. It also increases the depletion of the ozone layer, protecting the Earth from sunlight. Haze, eutrophication, and global temperature changes are some of the additional effects of air pollution[5].

Fuel utilization and industry are the primary sources of air pollution. When fossil fuels are burned, they release gases and compounds into the atmosphere. Furthermore, air pollution not only contributes to climate change but is exacerbated by a particularly damaging feedback mechanism[6]. The planet's temperature rises due to air pollution from carbon dioxide and methane. Smog occurs when the temperature is hotter and there is more UV radiation, and the heat increases this kind of air pollution. Climate change is boosting mold and pollen production (due to severe weather conditions and increased flooding) (due to a longer pollen season and more pollen production).

"Air pollution" is one of the most significant challenges of our era, not only because of the impact on climate change but as a result of increased mortality in both public and individual health. Pollution is described as the development [7] of chemicals which are dangerous to the environment for humans and other living things. Pollutants are hazardous, more than usual, solids, liquids, or gasses that harm the quality of our environment. The psychological effects of air pollution are enormous. Even in times of minimal air pollution, the health of individuals is vulnerable and delicate. Short-term exposure to air pollutants is all associated with COPD, cough,

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:03:26 UTC from IEEE Xplore. Restrictions apply.

shortness of breath, wheezing, asthma, respiratory disease, and a significant risk of hospitalization[8]. The long-term effects of air pollution include chronic disease, asthma, pulmonary insufficiency, and heart disease. In addition, air pollution seems to have several detrimental health repercussions, including psychiatric, respiratory, and chronic problems, for early human development[10].

Analysis of the air quality index as criteria ranging from 0 to 500. The higher the AQI value, the higher the levels of air pollution and the greater the risk of poor air quality. For example, air with an AQI of 50 or less is considered excellent quality, whereas air with an AQI of more than 300 is considered dangerous. Pollution is the biggest threat to human life, and it is worse in India than in any other country. The Air Quality Index (AQI) is a numeric scale used to measure and report air quality in a specific location on a particular day. PM10, PM2.5, NO2, SO2, CO, O3, NH3, and Pb are the eight major pollutants used to calculate the local AQI. As the AQI grows, so do the risks to public health.

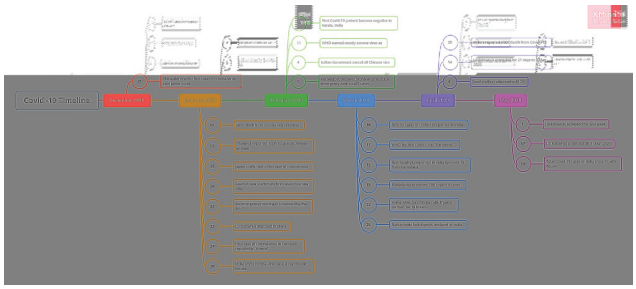


Fig. 1: COVID-19 timeline chart

II. Problem Statement

- To examine the relationships amongst toxic pollutants.
- The purpose of this research is to look at the impact of various environmental conditions on air quality.
- To determine the impact of altitude variations on air quality.
- To determine the impact of air quality on various regions.
- To determine the impact of various pollutants on air quality.
- Recognize different trends in the degree of air pollution throughout the course of a day.
- The purpose of this study is to see how the Covid-19 lockdown has affected air pollution levels in India, particularly in North India.

III. How does the AQI function

In India, AQI calculations include eight contaminants for which National Environmental Quality Standards (NAAQS) have been established. These pollutants are- PM2.5, PM10, N02, CO, SO2, O3, NH3, and lead.

- Individual pollutant sub-indices are derived from a monitoring station's 24-hour mean concentration value (8-hourly for CO and O3) and a set of health breakpoints. For this location, the AQI is the worst sub-index.
- Not all of the eight contaminants can be tested in every location. It is conceivable. Only AQI can be calculated if data for at least three pollutants, including PM2.5 or PM10, are available. In the absence of this information, the AQI estimate is deemed insufficient. A minimum of 16 hours of data is also required for the sub-index to be calculated.
- While there isn't enough data to calculate AQI, subsets of contaminants observed are calculated and distributed. The air quality status for each pollutant will be provided by the individual pollutant-wise sub-index.
- The AQI is provided in real-time through web-based technology. Without human intervention, the system collects data from continuous control stations and displays the Air Quality Index (AQI) based on average running values (e.g. AQI at 6 am on a day will incorporate data from 6 am on the previous day to the current day).
- An AQI calculator was created for manual stations, where data can be manually inputted to obtain an AQI value.

The overall AQI can only be calculated if data for at least three pollutants, either PM2.5 or PM10, are available. Similarly, a minimum of 16 hours of data is required for the calculation of the sub-index.

IV. Study area and Methodology

Our study is based on 26 cities of India which includes "Ahmedabad, Hyderabad, Patna, Gurugram, Amritsar, Mumbai, Chennai, Delhi, Bengaluru, Lucknow, Jorapokhar, Jaipur, Thiruvananthapuram, Amaravati, Brajrajnagar, talcher, Kolkata, Guwahati, Shillong, Chandigarh, Bhopal, Ernakulam, Kochi, Aizwal, Visakhapatnam, Coimbatore". The air quality index of 26 cities is studied from 01-01-2015 to 01-07-2020. The pollutant on which our studies is focused on PM2.5, PM10, N02, CO, SO2, O3, BTX.

India is officially known as the "Republic of India" is a country situated in South Asia. It is the second most populated country in the world, the seventh-biggest by geographic area, and the largest democracy.

The study was divided into two parts:

- Analysis during the years 2015-2020 of India's pollution levels:- It provides an outline of India's rising pollution levels and present situation.
- the consequences of the lockdown on India's pollution levels:- We'll look at pollution levels in India before and after the initial shutdown in this study. We'll compare pollution levels in the same months of 2019 to see whether there's a difference. We might also compare the levels of pollution in Northern India between now and the winter months of 2019. (The months of October and November)

The analysis of data in all these sections should provide us a realistic picture of Lockdown's impacts on pollution levels in India.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:03:26 UTC from IEEE Xplore. Restrictions apply.

TABLE I: AQI level and it's impact on health

AQI values	Condition	Colour code	Effect on public health
0-50	Good	Green	Minor repercussions.
51-100	Sastifactory	Light green	Sensitive person may experience minor breathing difficulties.
101-200	Moderate	Yellow	People with lungs, asthma and heart disorder have difficulty in breathing.
201-300	Poor	Orange	On prolonged exposure, most people experience breathing difficulties.
301-400	Very poor	Red	Long term exposure cause respiratory diseases.
401 and above	Severe	Maroon	Affect healthy persons as well as those who already have an illness.

TABLE II: AQI level for different pollutants in µg/m3.

AQI values	Color code	PM2.5 (for 24 hrs)	PM10 (for 24 hrs)	NO2 (for 24 hrs)	O3 (for 8 hrs)	CO (for 24 hrs)(mg/m3)	SO2 (for 24 hrs)	NH3 (for 24 hrs)	Pb (for 24 hrs)
0-50	Good	0-50	0-30	0-40	0-50	0-1	0-40	0-200	0-0.5
51-100	Satisfactory	51-100	31-60	41-80	51-100	1.1-2	41-80	201-400	0.6-1.0
101-200	Moderate	101-250	61-90	81-180	101-168	2.1-10	81-380	401-800	1.1-2.0
201-300	Poor	251-350	91-120	181-280	169-208	10.1-17	381-800	801-1200	2.1-3.0
301-400	Very poor	351-430	121-250	281-400	209-748	17.1-34	801-1600	1201-1800	3.1-3.5
401-500	Severe	430+	250+	400+	748+	34+	1600+	1800+	3.5+

V. Dataset used

For our study, we have used "Air Quality Data in India (2015-2020). It is an Air Quality Index (AQI), hourly data across the stations Cities in India. In this dataset, the total number of Cities is 26. Cities are "Ahmedabad, Hyderabad, Patna, Gurugram, Amritsar, Mumbai, Chennai, Delhi, Bengaluru, Lucknow, Jorapokhar, Jaipur, Thiruvananthapuram, Amaravati, Brajrajnagar, talcher, Kolkata, Guwahati, Shillong, Chandigarh, Bhopal, Ernakulam, Kochi, Aizwal Visakhapatnam, Coimbatore".

This dataset has:- City-data, City hour data, Station data, Station hour data, Station. In this, Station is controlled and monitored by the Central State agency which is the central pollution control board and state pollution control board respectively. This dataset consists of 15 parameters which are Date, PM2.5, PM10, NO, NO, NOX, NH3, CO, CO2, SO, O3, Benzene, Toluene, Xylene, AQI AQI Bucket.

A. Data pre-processing

- In this dataset, the date Column parameter is converted into the date-Time format.
- In this dataset, data is available between 01-01-2015 to 01-07-2020.
- After analyzing all the parameters of the dataset we have Combined Benzene, Toluene xylene Columns into one Column named at BTX.
- Similarly, we have combined PM2.5 PM10 columns into one as particulate matter.

VI. Simulation and Results

In our data, we have different types of pollutants. We have selected some eminent and distinguished pollutants ones which are PM2.5, PM10, N02, CO, SO2, O3, BTX.

Firstly, we have visualized and analyze the data yearly between 2015-2020 for PM2.5, PM10, N02, CO, SO2, O3, BTX and plotted a graph as shown in fig 2.

As we can see in fig 3, PM2.5 and PM10 decrease in 2020 as compared to 2019. Pollution concentration for PM2.5 and PM10 has a periodic influence having pollution concentration being greater in the winter months, particularly in comparison to the summer months. The level of SO2 began to climb after 2017, but there was a brief spike in 2015. The same trend may be seen at the BTX level as well. The O3 concentration is high between 2017-2019.

To analyze the level of pollution deeply and to understand the pattern, we have visualized and analyze the data monthly

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:03:26 UTC from IEEE Xplore. Restrictions apply.

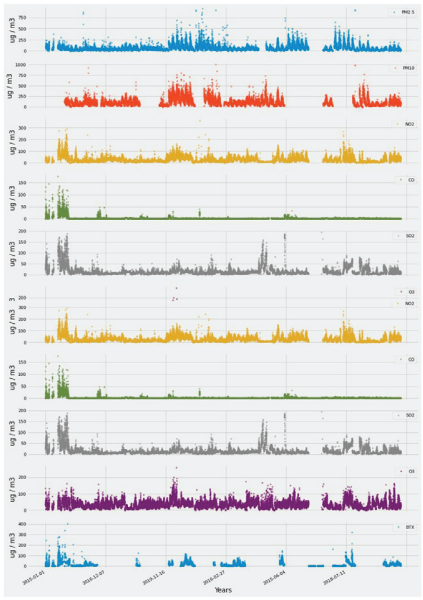


Fig. 2: Different pollutants level from 2015-2020

and yearly between 2015-2020 for PM2.5, PM10, N02, CO, SO2, O3, BTX and plotted a graph year-wise to understand the trend of pollutants and month-wise to understand the concentration of pollutants seasonally.

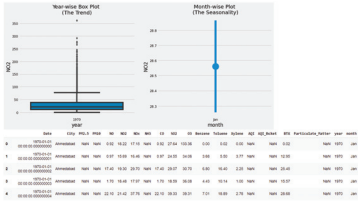


Fig. 3: NO2 level from 2015-2020

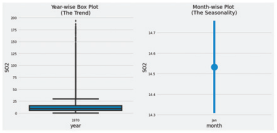


Fig. 4: SO2 level from 2015-2020

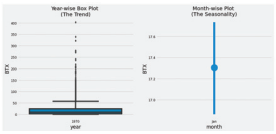


Fig. 5: PM2.5 level from 2015-2020

In fig 3,4,5,6,7 we can see, in the month of July and August, there is a noticeable downward trend in India's pollution level. It might be due to the fact that monsoon season arrived during these months. Interestingly, in April, the BTX levels indicate

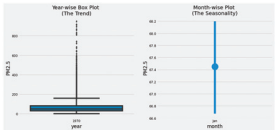


Fig. 6: PM10 level from 2015-2020

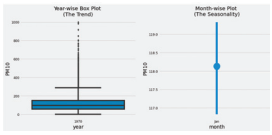


Fig. 7: BTX level from 2015-2020

a significant drop. After this, the environmental pollution concentration begins to rise, peaking during the winter months due to the burning of crop residuals, particularly in the northern parts, which is common throughout this season. After 2017, the amount of SO2 began to grow, albeit it had also risen suddenly in 2015. The same trends may be seen at the BTX level as well. The median reading for 2020 is usually lower than for previous years, indicating that pollution may have decreased recently.

After analyzing the trend of pollution year-wise and month-wise, now we have a look into the top 10 most pollutant cities from our dataset in each pollutant categories PM2.5, PM10, N02, CO, SO2, O3, BTX.

	City	PM2.5	City	PM10	City	NO2	City	SO2	City	CO	City	BTX
0	Patna	123.500000	0	Delhi	232.810000	0	Ahmedabad	58.180000	0	Ahmedabad	22.190000	0	Kolkata	36.230000
1	Delhi	117.200000	1	Gurugram	181.580000	1	Delhi	50.730000	1	Jorapokhar	33.650000	1	Lucknow	2.130000	1	Ahmedabad	37.110000
2	Gurugram	117.100000	2	Talcher	165.770000	2	Kolkata	40.400000	2	Talcher	28.490000	2	Delhi	1.980000	2	Delhi	35.960000
3	Lucknow	109.710000	3	Jorapokhar	149.660000	3	Patna	37.490000	3	Patna	22.130000	3	Talcher	1.850000	3	Patna	17.430000
4	Ahmedabad	67.850000	4	Patna	126.750000	4	Visakhapatnam	37.190000	4	Kochi	17.600000	4	Bengaluru	1.840000	4	Visakhapatnam	15.030000
5	Kolkata	64.360000	5	Brajrajnagar	124.220000	5	Lucknow	33.240000	5	Delhi	15.900000	5	Brajrajnagar	1.800000	5	Gurugram	14.600000
6	Jorapokhar	64.230000	6	Jaipur	123.480000	6	Jaipur	32.420000	6	Mumbai	15.200000	6	Ernakulam	1.630000	6	Amritsar	14.580000
7	Brajrajnagar	64.060000	7	Bhopal	119.320000	7	Bhopal	31.350000	7	Guwahati	14.660000	7	Patna	1.530000	7	Hyderabad	10.730000
8	Guwahati	63.690000	8	Guwahati	116.600000	8	Coimbatore	28.780000	8	Amaravati	14.260000	8	Kochi	1.300000	8	Chandigarh	9.090000
9	Talcher	61.410000	9	Kolkata	115.630000	9	Hyderabad	28.390000	9	Bhopal	13.060000	9	Gurugram	1.260000	9	Amaravati	3.680000

Fig. 8: Top pollutant cities from 2015-2020 on different pollutant

As we can see in fig 8, Patna has the highest concentration of PM2.5, Delhi has the highest concentration of PM10. Ahmedabad has the highest concentration of SO2 CO and Kolkata has the highest level of BTX(benzene, toluene, and xylene) concentration. To look at these values and understand them clearly, we have plotted a histogram as shown in fig 9 and 10.

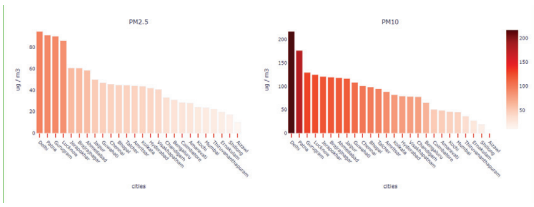


Fig. 9: 10 polluted cities in PM2.5 and PM10

Thirdly, we have understood the effect of lockdown on AQI in 6 cities:- Ahmedabad, Delhi, Mumbai, Chennai, Bengaluru, and Hyderabad. We will only look at data from 2019 onwards for this.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:03:26 UTC from IEEE Xplore. Restrictions apply.

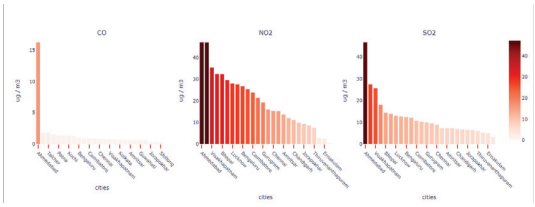


Fig. 10: 10 polluted cities in CO, NO2 and SO2

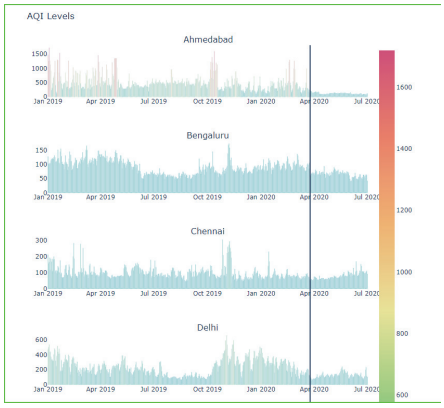


Fig. 11: Effect of lockdown on AQI in 6 cities:- Ahmedabad, Delhi, Mumbai, Chennai, Bengaluru, and Hyderabad from year 2019.

In fig 11,12 we can see the data on which India's first phase of lockdown went into force is indicated by black vertical lines. From January 2019 until the present, the graph above depicts the variance in various pollutant levels.

Unfortunately, all of the Indian cities listed above appear to have an extremely high level of pollution concentration. Undoubtedly, all of the cities under consideration look to be seeing a fast decrease after March 2020.

For better understanding, we have plotted gauge meter having labels:- Good, Satisfactory, Moderate, Poor, Very Poor, and Extreme representing AQI levels. For all the 3 cities mentioned above gauge- meter is shown in fig 13,14,15.

Lastly, to show the relationship between cities and type of pollutants(PM2.5, PM10, N02, CO, SO2, O3, BTX) we have

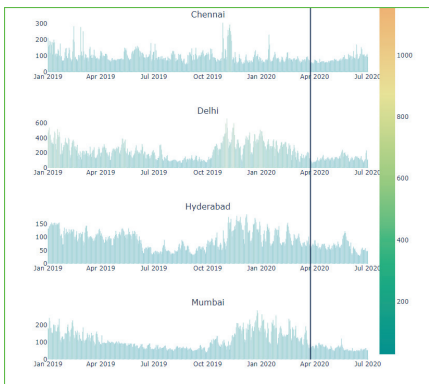


Fig. 12: Effect of lockdown on AQI in 6 cities:- Ahmedabad, Delhi, Mumbai, Chennai, Bengaluru, and Hyderabad from year 2019.

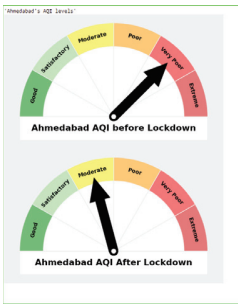


Fig. 13: Ahmedabad AQI before and after lockdown

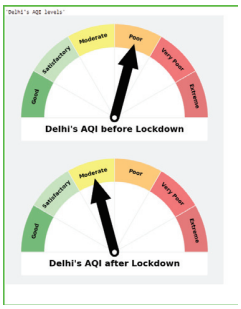


Fig. 14: Delhi AQI before and after lockdown

plotted heat map for the 2015-2020 year and for the 2020 year is plotted in fig 16 and 17 respectively.

VII. Conclusion

The Air Quality Index (AQI) has declined substantially throughout the country since Janta Curfew on 22 March to date. Summer pollution levels in India are lower than winter levels. Nonetheless, the present drop in pollution levels in the country is owing to the country's lockdown. As a result, all plants and other enterprises have been closed since the last week of March. The AQI scores in Indo-Gangetic cities improved dramatically, going from the higher to the lower end of the scale, with seven cities moving to the satisfactory category and three cities going to the 'good' category, respectively. Concentrations of PM10 and PM2,5 reduced by nearly half compared to the pre-lockdown period. NO2 and CO levels were also dramatically decreased during the lockdown. Air quality was improved by more than 60 percent in transport

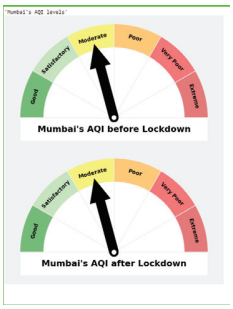


Fig. 15: Mumbai AQI before and after lockdown

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:03:26 UTC from IEEE Xplore. Restrictions apply.

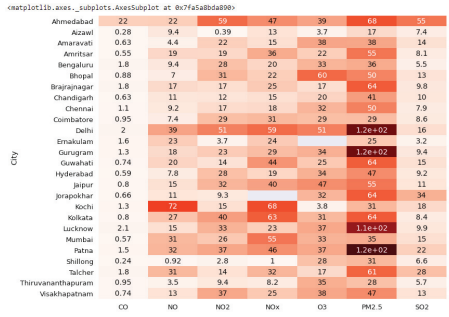


Fig. 16: Heat map for 2015-2020 year AQI

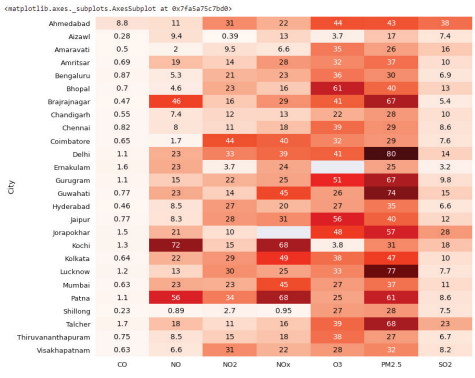


Fig. 17: Heat map for 2020 year AQI

and industrial regions. Air quality has improved considerably in central and eastern Delhi. On the second and fourth days of lock-down, the air quality improved by 40 to 50 percent.

Concentrations of PM10 and PM2,5 diminished by almost half contrasted with the pre-lockdown time frame. NO2 and CO levels were likewise significantly diminished during the lockdown. Air quality was improved by in excess of 60 percent in transport and mechanical districts. Air quality has improved significantly in focal and eastern Delhi. On the second and fourth long stretches of lock-down, the air quality improved by 40 to 50 percent.

During the lock-out period in 2020, Delhi, Mumbai, Hyderabad, Kolkata, and Chennai showed a considerable decrease in PM2.5 and AQI (Air Quality Index) as well as a downturn in the tropospheric NO2 level over the year. The air quality for the entire lock-up period has improved considerably, providing the municipal governments with critical information in drafting regulations and legislation to improve air quality. Delhi has also seen significant improvements, with high pollution levels being dealt with virtually all year. A good AQI is between 0 and 50, a decent AQI is between 51 and 100, a moderate AQI is between 101 and 200, a bad AQI is between 201 and 300, and an AQI of 401-500 is extremely poor. The Air Quality Index is a method of measuring air quality that takes several factors into consideration. Clean air is represented in the AQI. The lower the AQI, the better. Patna in Bihar and Delhi were the most polluted cities with a PM2.5 concentration of 123 and 117µg/m3 respectively. Delhi and Gurugram were the most polluted cities with a PM10 concentration of 232 and 191µg/m3 respectively. Ahmedabad has the highest level of NO2, SO2, and CO concentration of 59,55.2, and 22.1µg/m3 respectively. Nitrogen dioxide (NO2) is a greenhouse gas released by fossil-fuel burning and has been substantially decreased, averaging 13 percent in India and 31.5 percent in six towns. The national capital fell by 40 percent.

REFERENCES

[1] Amann M., Purohit P., Bhanarkar A.D., Bertok I., Borken-Kleefeld J., Cofala J., Heyes C., Kiesewetter G., Klimont Z., Liu J., Majumdar D. Managing future air quality in megacities: a case study for Delhi. Atmos. Environ. 2017;161:99–111.

[2] Anjum N.A. 2020. Good in the Worst: COVID-19 Restrictions and Ease in Global Air Pollution. Preprints. 2020040069. - DOI

[3] Bashir M.F., Ma B., Komal B., Bashir M.A., Tan D., Bashir M. Correlation between climate indicators and COVID-19 pandemic in New York, USA. Sci. Total Environ. 2020;728 - PMC - PubMed

[4] Beig G., Ghude S.D., Deshpande A. Indian Institute of Tropical Meteorology; 2010. Scientific Evaluation of Air Quality Standards and Defining Air Quality Index for India.

[5] Bezuglaya E.Y., Shchutskaya A.B., Smirnova I.V. Air pollution index and interpretation of measurements of toxic pollutant concentrations. Atmos. Environ. Part A. 1993;27(5):773–779.

[6] WHO (World Health Organization). World Health Organization Emergency Dashboard. https://covid19.who.int/region/searo/country/in. Accessed 8 July 2020 (2020)

[7] Hu, B., Guo, H., Zhou, P. Zheng, L. S. Characteristics of SARS-CoV-2 and COVID-19. Nat. Rev. Microbiol. https://doi.org/10.1038/s41579-020-00459-7 (2020).

[8] Chen LWA, Chien LC, Li Y, Lin G. Nonuniform impacts of COVID-19 lockdown on air quality over the United States. Sci. Total Environ. 2020 doi: 10.1016/j.scitotenv.2020.141105. - DOI - PMC - PubMed

[9] Berman JD, Ebisu K. Changes in US air pollution during the COVID-19 pandemic. Sci. Total Environ. 2020;739:139864. doi: 10.1016/j.scitotenv.2020.139864. - DOI - PMC - PubMed

[10] Unnithan, P. S. G. Kerala reports first Coronavirus case in India. India Today News (Jan 30, 2020), India. https://www.indiatoday.in/india/story/kerala-reports-first-confirmed-nov... (2020).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:03:26 UTC from IEEE Xplore. Restrictions apply.

Dear technology, you are not welcome: exploring ungratefulness towards technology

Naval Garg and Sarika Kumari

Abstract

Purpose – *Technology has become all-pervasive and one of the essential requirements of life. The blessings of science and technology are countless, and today, it is almost impossible to imagine life without applications of technology. And at the same time, positive psychologists have advocated practicing gratitude for the betterment of personal and professional growth and development. But ironically, the area of gratitude toward technology remains unexplored. The purpose of this study is to investigate the difference in gratitude for the technological and non-technological aspects of life.*

Design/methodology/approach – *The first study compared gratitude for three parameters – sense of abundance, social appreciation and appreciation for others. The second study was based on the "counting blessing" exercise, in which respondents were asked to write five things for which they feel grateful in life. The Friedman test and Wilcoxon signed-rank test were used to compare the ranked data. In the third study, respondents were distributed into two groups. Both groups were given a puzzle to solve; first, group members were informed that they would interact with other human beings, and the second group was informed that they would interact with bots.*

Findings – *All three studies concluded that respondents are relatively less grateful for the contributions of technologies.*

Originality/value – *This is probably the first study that investigated gratefulness for technology.*

Keywords *Ethics, Technology, Gratitude, Broaden-and-build theory, Human relations*

Paper type *Research paper*

Naval Garg is based at the Delhi Technological University, Delhi, India. Sarika Kumari is based at the Department of Telecommunications, New Delhi, India.

Introduction

> Gratitude is not only the greatest of virtues but the parent of all others. — Cicero
>
> Ingratitude! thou marble-hearted fiend, more hideous when thou show'st
>
> thee in a child than the sea-monster! — Shakespeare's King Lear

Contemporary researchers and practitioners have highlighted the two-pronged development in modern organizations. On the one hand, technology has become an indispensable part of organizational life. On the other hand, there is a growing inclination toward the institutionalization of human values, ethics and virtues at the workplace. These two complementary aspects of organizational growth aspire for a happy, comfortable, responsive and human workplace, which influences employees' psychological well-being and physical health. In the pursuit of ethical and values-based firms, modern managers bank upon several innovative and unconventional initiatives such as workplace spirituality, gratitude, constructive deviance, non-violent work behavior and organizational citizenship behavior.

Among these variables, institutionalizing an organizational culture based on gratitude was considered one of the most potent and newest tools to humanize the workforce. Positive psychologists have referred to gratitude as a moral virtue, an attitude, an emotion, a habit, a

Received 18 March 2021
Revised 14 April 2021
Accepted 5 November 2021

The authors confirm that the manuscript adheres to ethical guidelines. This includes requirements that research is conducted ethically, results are reported honestly, the submitted work is original and not (self-)plagiarized and authorship reflects individuals' contributions.
The author has no conflict of interest to declare.
We hereby declare that the present article is Compiled with Ethical Standards as below.
Disclosure of potential conflicts of interest – No conflict of interest.
Research involving human participants and/or animals – Informed consent of human respondents taken.

personality trait and a coping response (Emmons *et al.*, 2003). These different connotations of gratitude have two essential characteristics. First, the beneficiary recognizes the good done to him. This benefit may be a materialistic item like money, gift and promotion, or it may be a non-materialistic item such as happiness and satisfaction. Second, benefits should be attributed to some external agency. These external contributors may be human or non-human agencies like God, nature or good luck. Previous researchers have reported several benefits of gratitude, namely happiness, satisfaction and contentment, good physical and mental health, resilience, positive thinking and outlook, emotionality and spirituality, pro-social behavior and social intelligence. Besides ethical values like gratitude, modern organizations are also encouraging technology-enabled automation, artificial intelligence, deep learning, etc. The penetration and usage of these technologies have surged exponentially in recent times. Not only personal but also organizational life seems impossible without the application of modern science and technology. Researchers have highlighted encouraging inferences of technology-enabled responsive, prompt, comfortable and easy organizational processes and interactions.

Unfortunately, technology and gratitude are not studied simultaneously. And thus, it is necessary to research gratitude toward technology to achieve perfect complementaries of gratitude and technology. Previous studies have examined one's gratitude for diverse aspects of life like nature, fellow humans, God, amusing incidents, relatives or family members. However, probably no previous study has investigated one's gratefulness for all-pervasive technology. The present study tends to examine this unexplored domain with the help of three studies. The first study compares gratitude levels for technology and non-technology paradigms using a 16-statement questionnaire. The second study uses rank comparison to investigate the gratitude difference between the two groups. The third study examines the difference in the experience of gratefulness in two experimental settings.

Theoretical framework

Researchers have elaborated on several different definitions, concepts and constructs of gratitude. McCullough *et al.* (2002) defined gratitude "as a generalized tendency to recognize and respond with grateful emotion to the roles of other people's benevolence in the positive experiences and outcomes that one obtains" (p. 112). This conceptualization of gratitude focused on "benefit-triggered gratitude," wherein gratitude is experienced only when one recognizes and attributes benefits to another person (McCullough *et al.*, 2001). Emmons and McCullough (2003) elaborated a broader connotation of gratitude as "an emotion, an attitude, a moral virtue, a habit, a personality trait, or a coping response" (p. 377). This broader conceptualization inscribed common tendency, trait, habit or virtue of people to appreciate pleasant things of life and positive emotions and responses to particular situations or reflections, that is, felt gratitude (Wood *et al.*, 2008). Thus, Emmons and McCullough (2003) discussed the cognitive and emotional components of gratitude. Watkins *et al.* (2003) further broadened the concept stating that gratitude is "the affective trait of gratitude may be thought of as a predisposition to experience gratitude" (p. 432). Watkins *et al.* (2003) argued that researchers must explore gratitude as both "state" and "trait." Gratitude state could be defined as "a feeling of thankful appreciation for favors received" (Guralnik, 1971, p. 327, as cited in Watkins *et al.*, 2003). And trait gratitude is "a predisposition to experience the state of gratitude" (Watkins *et al.*, 2003, p. 432).

Watkins *et al.* (2003) acknowledged three critical characteristics of a grateful individual. First, a grateful person always remains in the sense of abundance or lack of sense of deprivation. Such individuals count, acknowledge and honor all physical, material and psychological assets that one possesses. Second, a grateful individual appreciates other's contributions to one's life. And finally, a grateful person enjoys the simplest pleasures of life, such as a pleasant meal, excellent weather and impromptu meeting with an old friend. These simple incidents are pretty familiar to all, but only a grateful person could

appreciate these mundane life events. Adler and Fagley (2005) interpretation of gratitude included thankfulness for benefits received from non-human entities like nature or God (e.g. being thankful for a good and deep sleep at night). Friedman (2000) proposed an exhaustive list of things for which one experiences gratefulness. The list includes:

- people, condition and state of life;
- all that one learns, experiences and receives;
- cognitive, moral and spiritual resources;
- one's contribution to others and the society; and
- prosperity, abundance and blessings.

Further, Garg (2017a, 2017b) introduced "kritagayata" as a synonym of gratitude and recommended that an actual grateful person will remain thankful for even unpleasant and unexpected incidents of life. Such a person will appreciate the chance of growth and development in a most tiring and challenging phase of life (Saxena *et al.*, 2020).

Several researchers and practitioners reported positive individual, group and organizational benefits of gratitude. Scholars reported linkage between gratefulness and different organizational variables like relationship satisfaction and contentment (Wood *et al.*, 2010; Algoe *et al.*, 2010; Kubacka *et al.*, 2011), good health (Wood *et al.*, 2010; Emmons and McCullough, 2003), resilience (Wood *et al.*, 2010), lower level of stress and anxiety (Wood *et al.*, 2010), positive thinking and outlook (Wood *et al.*, 2009), emotionality and spirituality (McCullough *et al.*, 2002), pro-social behavior (Bartlett and DeSteno, 2006; Andersson *et al.*, 2007), social intelligence (Garg and Gera, 2019), leadership (Garg, 2020), life satisfaction (Emmons and McCullough, 2003) and job satisfaction (Waters, 2012). Also, several researchers, including Emmons and McCullough (2003), Watkins *et al.* (2003), Froh *et al.* (2014) and Jones *et al.* (2013), revealed empirical evidence of a direct relationship between gratitude and happiness.

These associations are theoretically explained with the help of Fredrickson's (2004) broaden-and-build theory of positive emotions. She elaborated that positive emotions like gratitude "broaden people's momentary thought-action repertoires and build their enduring social resources" (Fredrickson, 2004, p. 147). In simple words, gratitude may bring out a well-thought response from the beneficiary instead of a tit-for-tat reaction. A grateful person explores various responses more empathetically and innovatively. Such reflections of gratitude by the receiver need to be intended to benefit the giver only. But the receiver may aspire to return benefit to the third party, including fellow human beings, nature and society at large (Jans-Beken *et al.*, 2019). The author suggested that gratitude may widen thought–action tendencies in three significant ways: expanding the scope of attention, the range of cognition and the scope of action. Several other authors, including McCullough *et al.* (2001) and Watkins *et al.* (2014), appreciated the broaden-and-build perspective of gratitude, stating that it can aspire beneficiary to realize and acknowledge others' good deed in its favor. Previous researchers have suggested that grateful people are more attentive and appreciative of others' acts of goodwill.

Later, Fredrickson elaborated that gratitude builds resources for a beneficiary in several ways. It develops an empathetic and compassionate relationship between the giver and receiver, and thus, it promotes a cohesive bond between two parties of exchange. This extraordinary and individualized bond may facilitate a lifelong association between two parties. Also, gratitude expands the receiver's "more general skills for loving" (Fredrickson, 2004, p. 152). A grateful person focuses on the positive and blessing part of a natural or human-made incident. Such a person radiates love, affection and caring for others (Yang *et al.*, 2018; Salces-Cubero *et al.*, 2018). This expanded thought–action repertoires and

loving bond make recipients value others' contributions in their life. And in cases where benefits could not be attributed to human factors like nature, a person's gratitude reflects in the form of reasonable use of natural resources. Countless researchers have reported the positive impacts of nature on human beings. These researchers include Bratman *et al.* (2015), Clay (2001), Nichols (2016), Taylor *et al.* (2002) and Weinstein *et al.* (2009). The realization of selfless blessings of nature develops a deep sense of gratitude. This gratuitous person tends to reciprocate with care and concern for mother nature.

The same theoretical foundation explains the rationale of the current study. Probably, no study ever explored man's gratitude toward technology. Today, technology has become all-pervasive and one of the essential requirements of life. The blessings of science and technology are countless, and today, it is almost impossible to imagine life without applications of technology. Amid this indispensability of technology, it is imperative to examine gratitude for technology for various reasons. First, following the broaden-and-build theory, gratitude for technology may result in increased awareness and resilience against the misuse of technology. The broadened "thought-action repertoires" and development of a bond based on affection and care (just like nature) may entice a grateful person to shun technological abuse and to adopt the innovative and productive usage of modern technology. The argument finds theoretical support from Watkins's (2014) amplification theory of gratitude also. Watkins argued that as an amplifier increases the power of a signal, gratitude amplifies the message of goodness in one's life. "Thus, by experiencing and expressing gratitude, information on the goodness in one's life is strengthened" (Watkins, 2014, p. 252). Using an analogy from the amplification theory, gratitude toward technology is bound to intensify the beneficial usage of technology, and subsequently, the misuse of technology may reduce.

Second, the realization of gratitude toward technology stimulates the beneficiary to revert acts of goodness. As technology is a non-human benefactor, the receiver tends to think differently to pay off the debt of technology. And one of the probable and doable actions is efforts to reduce the digital divide. It may help in the proliferation of technology to the downtrodden and needy people of the society. Benefactors may ponder upon various other acts of indebtedness, such as judicious use of technology, gifting technological knowhow and proper treatment of digital wastes. Further, a few studies like Ghandeharioun *et al.* (2016) and Godawa (2019) discussed "how technology enables the formation of gratuitous attitude." These authors advocated that electronic devices like Emails and Apps may help in practicing several gratitude interventions such as blessing count, daily gratitude diary and express gratitude openly. However, these studies and other similar endeavor to explore technology–gratitude interfaces are incomplete without an adequate understanding of gratitude toward technology. Gratitude studies are also incomplete without adding technology to their subject matter. Gratitude literature has so far missed the inclusion of technology, which is an essential part of human life. Such an addition will surely enrich the domain of gratitude.

Based on the above-discussed premises, the present study tends to investigate gratitude toward technology with the help of the following three studies.

Study 1

The objective of Study 1 was to compare gratefulness toward the technological and non-technological aspects of human life. Three dimensions of gratitude – *Lack of sense of deprivation* (*LOSD*), *appreciation of others* (*AO*) and *simple appreciation* (*SO*) – were compared. The sample comprised 180 students enrolled in the MBA program of state-level university. These students were randomly assigned to two groups. Respondents of Group 1 were given Watkins *et al.* (2003) Gratitude, Resentment and Appreciation Scale (GRAT). The same scale was adjusted to measure gratitude toward technology. And the gratitude of respondents of the second group for technology was accessed using this revised scale. The gratitude scale consisted of 16 items to capture three dimensions of gratitude – LOSD

(six statements), AO (four statements) and SO (six statements) – on a nine-point scale ranging from 1 (strongly disagree) to 9 (strongly agree). Students were approached directly with the help of a structured questionnaire comprising of GRAT-16 and demographic-information-related information such as age, gender, specialization and working experience. The sample consisted of 121 male (68%) and 59 (32%) female respondents. The mean age of students was 21.7 years. Nearly 30% of respondents reported working experience. As per previous researchers, when the population size is unknown, the minimum sample size should be at least ten times the number of statements. As statements in GRAT-16 are 16, a sample of more than 176 is desired to achieve a ratio of 11:1.

Descriptive statistics and correlation matrix of both groups were reported in Tables 1 and 2, respectively. The reliability of the measure was accessed using Cronbach's alpha, which represents the mean of all possible split-half correlations for a set of items (internal consistency). For Group 1, all values of Cronbach's alpha are higher than 0.60, which is acceptable for measures having few statements. The same is true for Group 2 too. Later, Cronbach and Shavelson (2004) advocated measuring composite reliability along with Cronbach's alpha. The value of composite reliability should be greater than 0.70 to assure reliability (Hair *et al.*, 2010). Further, convergent validity is ensured using average variance explained, whose value should be greater than 0.50 (Fornell and Larcker, 1981). Variance inflation factor of less than 5 negates the issue of multicollinearity (Kline, 2009). The pre-validity scale (GRAT-16) was used to deal with common method variance. Also, respondents were promised confidentiality of their information. When all items were subjected to principal component analysis, no single factor explained major variations in the measure. It negates the possibility of common method variance according to Harman's one-factor approach.

Tables 3 and 4 elaborated mean and standard deviation values of all 16 statements of GRAT-16. It was reported that mean values were higher for Group 1 for all statements. It signified that respondents' gratitude was less for technology than for other things of life. Table 5 investigated the statistical significance of the difference of gratefulness between Groups 1 and 2. Mean values of three dimensions of gratitude (sense of abundance, SO and AO) are higher for Group 1. And *t*-vales concluded that respondents of Group 1 were statistically significantly more grateful than respondents of Group 2.

Table 1 Descriptive statistics and correlation matrix (Group 1)

Dimension of gratitude	*1*	*2*	*3*	α	*CR*	*AVE*	*VIF*
Sense of abundance	1			0.628	0.818	0.674	2.78
Appreciation for others	0.845**	0.882**	1	0.610	0.732	0.680	1.95
Simple appreciation	0.804**	1		0.614	0.726	0.643	2.91

Sources: Primary data, **Correlation is significant at the 0.01 level (two-tailed), α – Cronbach's alpha, CR – composite reliability, AVE – average variance extracted; VIF – variance inflation index

Table 2 Descriptive statistics and correlation matrix (Group 2)

Dimension of gratitude	*1*	*2*	*3*	*A*	*CR*	*AVE*	*VIF*
Sense of abundance	1			0.748	0.835	0.793	2.34
Appreciation for others	0.474**	0.517**	1	0.774	0.879	0.730	3.46
Simple appreciation	0.628**	1		0.762	0.847	0.777	1.96

Sources: Primary data, **Correlation is significant at the 0.01 level (two-tailed), α – Cronbach's alpha, CR – composite reliability, AVE – average variance extracted; VIF – variance inflation index

Table 3 Descriptive statistics (Group 1)

Serial number	*Statement*	*Mean*	*SD*
1	I could not have gotten where I am today without the help of many people	5.91	2.86
2	Life has been good to me	5.72	2.30
3	There never seems to be enough to go around, and I never seem to get my share	5.81	1.47
4	Oftentimes, I have been overwhelmed at the beauty of nature	5.97	2.12
5	Although I think it is important to feel good about your accomplishments, I think that it is also important to remember how others have contributed to my accomplishments	6.03	2.53
6	I really do not think that I have gotten all the good things that I deserve in life	5.24	1.55
7	Every Fall, I really enjoy watching the leaves change colors	5.47	2.07
8	Although I am basically in control of my life, I cannot help but think about all those who have supported me and helped me along the way	5.84	2.23
9	I think that it is important to "Stop and smell the roses."	5.64	2.17
10	More bad things have happened to me in my life than I deserve	5.49	1.58
11	Because of what I have gone through in my life, I really feel like the world owes me something	5.53	1.57
12	I think that it is important to pause often to "count my blessings"	5.66	2.07
13	I think it is important to enjoy the simple things in life	6.38	2.31
14	I feel deeply appreciative for the things others have done for me in my life	6.12	2.41
15	For some reason I do not seem to get the advantages that others get	5.86	1.76
16	I think it is important to appreciate each day that I am alive	6.46	2.28

Source: Primary data

Study 2

In Study 2, gratitude toward technology and other aspects of life was compared with the help of a gratitude exercise – "Blessing Counting." This study was conducted in two phases. In the first phase, 50 students were asked to write down five things of life they feel grateful for. The collected information is tabulated and shown in Table 6. The results revealed that respondents were most grateful for parents (96%), physical and mental health (78%), life (68%), friend (60%) and financial security (42%). Only 18% of respondents were grateful for technology. It reflected that gratefulness for technology was relatively lower than other aspects of life. In the second phase of the study, the top four grateful items, that is, parents, health, life, friend and technology, were taken. Another set of 50 students were asked to rank these five items from "most grateful" (1) to "the least grateful" (5). As five items were chosen for future analysis, a sample of 50 was desired as per the thumb rule of "ten multiplied by the number of statements/items." This ranking-based data was subjected to the Friedman test, which is a non-parametric test used to examine the differences between items measured on an ordinal scale. Table 7 elaborated the mean rank of five items – parents, health, life, friend and technology. The mean rank for different items were 1.59 for parents, 2.42 for health, 2.60 for life, 3.80 for friend and 4.59 for technology. It was also observed that the difference in gratitude was statistically significant for five measured items with $\chi^2(4) = 113.392$ and $p = 0.000$. Although it was evident that respondents were significantly least grateful for technology, the Wilcoxon signed-rank test was also applied to explore the difference between two items independently, that is, between parents and technology, health and technology, life and technology and friend and technology. Table 8 indicated the results of the Wilcoxon signed-rank test. The results suggested that gratitude toward technology was significantly lower than other items of the study.

Table 4 Descriptive statistics (Group 2)

Serial number	*Statement*	*Mean*	*SD*
1	I could not have gotten where I am today without the help of technology	3.04	1.13
2	Experience of technology has been good to me	3.26	1.22
3	I could not realize all benefits of the technological devices	4.26	1.25
4	Oftentimes I have been overwhelmed at the contribution of technology	3.73	1.35
5	Although I think it is important to feel good about your accomplishments, I think that it is also important to remember how technologies have contributed to my accomplishments	3.44	1.40
6	I really do not think that I have cherished all the benefits of technology	4.14	1.26
7	I really enjoy new version/features of mobile phones	3.82	1.30
8	Although I am basically in control of my life, I cannot help but think about different technological inventions and devices who have supported me and helped me along the way	3.60	1.31
9	I think that it is important to honor presence of technology in every sphere of life	3.57	1.38
10	Technology has resulted into more bad things to me in my life than I deserve	4.09	1.23
11	Because of the contribution of technology in my life, I feel like reciprocating the benefits to others	4.29	1.24
12	I think that it is important to pause often to "count blessings of technology"	3.66	1.36
13	I think it is important to enjoy simple electronic devices	3.17	1.11
14	I feel deeply appreciative for the things technology have done for me in my life	3.30	1.25
15	For some reason, I do not seem to get the advantages of technology that others get	4.46	1.31
16	I think it is important to appreciate each day that I live in a technologically advanced society	3.03	1.06

Source: Primary data

Table 5 Variation in gratitude (Group 1 vs Group 2)

Dimension of gratitude	*Mean (Group 1)*	*Mean (Group 2)*	t-*value*	*Significance*
Sense of abundance	5.78	3.64	12.39	0.000*
Appreciation for others	6.20	3.48	12.36	000*
Simple appreciation	5.60	3.83	10.96	000*

Sources: Primary data; *significant at 0.05 level of significance

Table 6 Results of blessing counting exercise

Serial number	*Item*	*No. of students*	*(%)*
1	Parents	48	96
2	Health – Physical and mental	41	78
3	Life	34	68
4	Friend	30	60
5	Financial status	21	42
6	Technology	09	18
7	Country	10	20
8	Flora and fauna	12	24
9	Love of life (boyfriend or girlfriend)	19	38
10	Miscellaneous*	26	52

Sources: Primary data; *Item count less than ten were included in Miscellaneous (except technology)

Table 7 Results of the Friedman test

Serial number	*Item*	*Mean rank*	*Chi-square*	*Significance*
1	Parents	1.59	113.392	0.000*
2	Health – Physical and mental	2.42	(Degree of freedom - 4)	
3	Life	2.60		
4	Friend	3.80		
5	Technology	4.59		

Sources: Primary data; *significant at 0.05 level of significance

Table 8 Results of Wilcoxon signed-rank test

Serial number	*Interaction*	*Z-value*	*Significance*
1	Parents–Technology	6.216	0.000*
2	Health–Technology	5.190	0.000*
3	Life–Technology	5.122	0.000*
4	Friend–Technology	4.317	0.000*

Sources: Primary data; *significant at 0.05 level of significance

Study 3

The study aspired to examine the gratitude difference with the help of experimental design. The students were asked to register voluntarily to participate in the experiment. In all, 80 students registered for the research, and they were randomly distributed into two groups. Group 1 comprised 22 male and 18 female students. And Group 2 had 24 boys and 16 girls. The mean ages of both groups were 19.7 and 20.3 years, respectively. These groups consisted of students of B.Tech, BBA, BA and MBA. The sample size of 40 (for each group) is based on the thumb rule of "ten times of statements." As statements are four, a sample of at least 40 is required. Group 1 respondents were told that they would interact with another human partner to solve a given puzzle. Group 2 students were told that they would interact with a bot to solve the same problem, but in reality, they communicated with a human partner only. The research design created an artificial system where respondents felt that they were interacting with a bot while they were in contact with another respondent only. Steinfeld *et al.* (2009) suggested that such an experimental design permits the moderation of interaction to stimulate human and machine interaction. Thus, 20 dyads in both groups were created. These respondents interacted through one of the most commonly used online platform – Whatsapp – for its simplicity and familiarity to all participants. After the experiment, participants were requested to respond to the Simple Appreciation Scale of GRAT-16. The language of the scale was suitably adjusted to accommodate the experimental conditions. Table 9 summarizes the results of the experiments. It was reported that mean values were greater for Group 1 for all four statements and overall SO. And *t*-values reflected that the variations in mean values were statistically significant. In a nutshell, Group 1, where respondents interacted with a human partner, was reported to be more grateful for their partners than Group 2, where members perceived to have interacted with a bot. In conclusion, all three studies concluded that respondents were relatively less thankful for technology.

Findings/results

The present study is one of the pioneer studies that investigated gratitude toward the technological aspect of human life. Banking upon broaden-and-build and amplification theory, three studies were conducted to compare gratitude for technological and non-technological elements of life. Study 1 used the GRAT-16 scale for the same. Gratitude for technology was reported to be significantly lesser for all 16

Table 9 Descriptive statistics and gratitude variation (Group 1 vs Group 2)

Serial number	Statement	Mean G-1	Mean G-2	SD G-1	SD G-2	t-value	Significance
1	I could not have solved the problem without the help of my team member/technology	6.20	3.97	2.58	1.18	5.016	0.00*
2	Although I think it is important to feel good about my accomplishments, I think that it is also important to remember how my fellow/technology have contributed to my accomplishments	5.92	4.05	2.57	1.19	4.521	0.00*
3	Although I could have solved the problem, I think my team member/technology supported me and helped me along the way	5.67	3.57	1.77	1.39	5.842	0.00*
4	I feel deeply appreciative for the things my fellow member/ technology have done for me to solve the given problem	6.17	3.25	2.55	1.23	6.430	0.00*
	Simple appreciation for Team member/Technology	5.99	3.71	2.06	0.82	6.631	0.00*

Sources: Primary data; *significant at 0.05 level of significance

statements of gratitude scale. The first study also concluded that the mean value of the three dimensions of gratitude – a sense of abundance (mean for technology: 3.64 and mean for non-technology: 5.78), AO (mean for technology: 3.48 and mean for non-technology: 6.20) and SO (mean for technology: 3.83 and mean for non-technology: 5.60). Gratitude for technology was reported highest for the SO of technology (mean = 3.83) and lowest for AO (mean = 3.48). However, general gratitude was observed greatest for AO (mean = 6.20) and lowest for SO (mean = 5.60). Thus, the ranking order of dimensions of gratitude is different for Groups 1 and 2 of the study. The finding bears significance in the selection of appropriate gratitude interventions to elicit gratitude toward technology. A general exercise of counting the contributions of technology in one's life is sufficient to cherish a simple appreciation gratitude state.

Study 2 was conducted in two stages. The results of the first stage highlighted that only 18% of respondents were grateful for technology in comparison to 96% for parents, 78% for health, 68% for life and 60% for a friend. These results are quite annoying, given the level and degree of penetration of technology in human life. It signifies that technology has not been appreciated enough for its contribution. One cannot deny the utility of technology for parents. Pre-conception processes like IVF, IUI and machines such as Ultrasound, Computed Tomography scan (CT scan) during pregnancy and surgical procedures like Cesarean, vacations of toddlers and genetic disorder tests after delivery have become a regular part of one's life. A truly grateful person also needs to realize these gifts of technology. In the second stage of Study 2, gratitude toward technology was ranked lowest among five items of gratefulness. Study 3 uses an experimental design to examine the gratefulness of respondents for a partner or bot for their help in solving a given puzzle. The results of the third study also corroborate that of the first two studies. After the experiment, students report a significantly higher level of gratefulness for a human partner than a bot partner. In a nutshell, technology has not been accredited its due credit for its immense contribution to human life.

Discussion and conclusion

The study makes some useful theoretical and practical contributions. It enriches the literature of gratitude in two possible ways. First, as discussed earlier, this is probably the first study that tried to examine gratefulness toward technology. Using a robust, rigor and comprehensive methodology highlighted the unfortunate plight of gratitude toward technology. Second, the present literature has not accounted for the technological aspect

of gratitude. No theory or scale is available that could explain the intricacies of gratitude–technology dynamics. Lack of studies in the field restricts the comparison of results of the present study with previously claimed facts and findings. The study calls for more researchers to examine this neglected sub-theme of gratitude study. The results of the current study could also be encouraging for experts working in the domain of human–computer interface. In the age of artificial intelligence and emotional intelligence, virtue as crucial as gratitude cannot be overlooked. The automated algorithmic-based decision platform tends to inscribe an ethical component in decision-making models. Researchers can use scientifically accumulated knowledge of gratitude for ethical decision-making. Researchers are also experimenting with several innovative interfaces of human–computer collaboration, and here, a gratuitous user may prove to be an appropriate collaborator.

Diverse theories of gratitude supported thought–action repertoire, which may help in the appropriate usage of modern technologies. Although these theoretical explanations require sound empirical evidence, contemporary organizations' managers could use different gratitude interventions to develop gratitude for technology among their employees. Technology and technology-enabled devices are inseparable parts of modern companies. An experience of gratefulness toward technological advancement will make employees realize the importance of machines in their personal and professional life. And one of the expressions of thankfulness for technology is appropriate, timely and reasonable use of modern tools and techniques. It may cut downtime and resource wastages for organizations. Also, the experience of gratitude toward technology may propel similar feelings for colleagues, which leads to trustworthy and cordial relations among employees. Based on organizational culture and other environmental constraints, one may select a combination of varying gratitude practices and interventions. These interventions include "gratitude journals" (Froh *et al.*, 2008, 2007), "gratitude letters" (Seligman, 2002; Sergeant and Mongrain, 2011), "gratitude letters visits," "cultivating mindfulness" (Langer, 2014; Langer *et al.*, 2010), "gratitude collages" (Reivich, 2009), "sharing positive events" (Gable *et al.*, 2010), "acts of kindness" (Sheldon *et al.*, 2012) and "meditation and progressive muscle relaxation techniques" (Shapiro *et al.*, 2005). Leadership also plays a vital role in the institutionalization of the state of gratitude for technology. Positive emotions like gratitude are contagious (Watkins, 2004), and gratuitous behavior by top management is doubly infectious (Garg, 2020). Thus, if a manager exhibits thankfulness for technology, then the subordinates are also likely to follow suit.

The paper also hints toward a probable amalgamation of two distinct organizational paradigms – human values and technology. Future organizations require an astute combination of values and technology to satisfy human wishes and demands. Human values like gratitude, compassion and love are necessary to prevent the mechanization of human efforts. And technology ensures vices-free human actions. Technology not only provides comfort and speed to human efforts but also keeps activities free from nepotism, favortism, stereotyping, halo and horn effect. An experience of gratitude also aspires to increase the acceptability of new technology.

The contributions of the current study have come in the light of few limitations. First, all three studies were based on a cross-sectional research design. Also, relevant pieces of information were collected with the help of self-reported measures. Future researchers may use a longitudinal research design. Also, the use of double-source data and replication of the study in different cultural settings are required for the generalization of results. Second, the present study was conducted on students (post-millennial generation) with little or no work experience. It is a well-known fact that ethical orientation and technology adaptability varies across different generations. It would be interesting to conduct this study in an organizational context comprising of people belonging to all three generations, that is, post-millennial, millennial and pre-millennial. Further, this study did not examine the probable reasons for the low level of gratefulness for technology. Future researchers may ponder upon several individuals, groups and organizational factors affecting gratitude toward technology.

References

Adler, M.G. and Fagley, N.S. (2005), "Appreciation: individual differences in finding value and meaning as a unique predictor of subjective well-being", *Journal of Personality*, Vol. 73 No. 1, pp. 79-114.

Andersson, L.M., Giacalone, R.A. and Jurkiewicz, C.L. (2006), "On the relationship of hope and gratitude to corporate social responsibility", *Journal of Business Ethics*, Vol. 70 No. 4, pp. 401-409, doi: 10.1007/s10551-006-9118-1.

Bartlett, M. and Desteno, D. (2006), "Gratitude and prosocial behavior helping when it costs you", *Psychological Science*, Vol. 17 No. 4, pp. 319-325, doi: 10.1111/j.1467-9280.2006.01705.x.

Bratman, G.N., Daily, G.C., Levy, B.J. and Gross, J.J. (2015), "The benefits of nature experience: improved affect and cognition", *Landscape and Urban Planning*, Vol. 138, pp. 41-50.

Clay, R.A. (2001), "Green is good for you", *American Psychology Association*, Vol. 32 No. 4, pp. 40-50.

Cronbach, L.J. and Shavelson, R.J. (2004), "My current thoughts on coefficient alpha and successor procedures", *Educational and Psychological Measurement*, Vol. 64 No. 3, pp. 391-418, doi: 10. 1177/0013164404266386.

Emmons, R.A. and McCullough, M.E. (2003), "Counting blessings versus burdens: an experimental investigation of gratitude and subjective well-being in daily life", *Journal of Personality and Social Psychology*, Vol. 84 No. 2, pp. 377-389, doi: 10.1037/0022-3514. 84.2.377.

Emmons, R.A., McCullough, M.E. and Tsang, J.-A. (2003), "The assessment of gratitude", in Lopez, S.J. and Snyder, C.R. (Eds), *Positive Psychological Assessment: A Handbook of Models and Measures*, American Psychological Association, pp. 327-341, doi: 10.1037/10612-021.

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, doi: 10.1177/002224378101800104.

Friedman, H. (2000), "Long-term relations of personality and health: dynamisms, mechanisms, tropisms", *Journal of Personality*, Vol. 68 No. 1, pp. 1089-1107, doi: 10.1111/1467-6494.00127.

Froh, J.J., Bono, G., Fan, J., Emmons, R.A., Henderson, K., Harris, C. and Leggio, H. (2014), "Nice thinking! An educational intervention that teaches children to think gratefully", *School Psychology Review*, Vol. 43 No. 2, pp. 132-152.

Froh, J.J., Miller, D.N. and Snyder, S.F. (2007), "Gratitude in children and adolescents: development, assessment, and school-based intervention", *School Psychology Forum*, Vol. 2 No. 1, pp. 1-13.

Froh, J.J., Sefick, W.J. and Emmons, R.A. (2008), "Counting blessings in early adolescents: an experimental study of gratitude and subjective well-being", *Journal of School Psychology*, Vol. 46 No. 2, pp. 213-233.

Gable, S.L., Reis, H.T., Impett, E.A. and Asher, E.R. (2010), "What do you do when things go right? The intrapersonal and interpersonal benefits of sharing positive events", *Journal of Personality and Social Psychology*, Vol. 87 No. 2, pp. 228-245.

Garg, N. (2017a), "Workplace spirituality and organizational commitment: an empirical exploration", *Purushartha-A Journal of Management, Ethics, and Spirituality*, Vol. 10 No. 2, pp. 8-24.

Garg, N. (2017b), "Workplace spirituality and organizational performance in Indian context: mediating effect of organizational commitment, work motivation, and employee engagement", *South Asian Journal of Human Resources Management*, Vol. 4 No. 2, pp. 191-211.

Garg, N. (2020), "Exploring role of gratitude in developing teacher leadership in Indian universities", *International Journal of Educational Management*, Vol. 34 No. 5, pp. 881-901, doi: 10.1108/ijem-07- 2019-0253.

Garg, N. and Gera, S. (2020), "Gratitude and leadership in higher education institutions: exploring the mediating role of social intelligence among teachers", *Journal of Applied Research in Higher Education*, Vol. 12 No. 5, pp. 915-926, doi: 10.1108/JARHE-09-2019-0241.

Ghandeharioun, A., Azaria, A., Taylor, S. and Picard, R.W. (2016), "'Kind and Grateful': a context-sensitive smartphone app utilizing inspirational content to promote gratitude", *Psych Well-Being*, Vol. 6 No. 9, pp. 1-21, doi: 10.1186/s13612-016-0046-2.

Guralnik, D.B. (1971), *Webster's New World Dictionary*, Southwestern Co.

Hair, J., Black, W., Babin, B. and Anderson, R. (2009), *Multivariate Data Analysis*, Prentice Hall, London.

Kubacka, K., Finkenauer, C., Rusbult, C. and Keijsers, L. (2011), "Maintaining close relationships", *Personality & Social Psychology Bulletin*, Vol. 37 No. 10, pp. 1362-1375, doi: 10.1177/0146167211412196.

Jans-Beken, L., Jacobs, N., Janssens, M., Peeters, S., Reijnders, J., Lechner, L. and Lataster, J. (2019), "Gratitude and health: an updated review", *The Journal of Positive Psychology*, Vol. 15 No. 6, pp. 743-782, doi: 10.1080/17439760.2019.1651888.

Jones, C.N., You, S. and Furlong, M.J. (2013), "A preliminary examination of covitality as integrated well-being in college students", *Social Indicators Research*, Vol. 111 No. 2, pp. 511-526.

Langer, E.J. (2014), *Mindfulness Forward and Back*, Wiley Online Library, doi: 10.1002/9781118294895.

Langer, E., Djikic, M., Pirson, M., Madenci, A. and Donohue, R. (2010), "Believing is seeing: using mindlessness (mindfully) to improve visual acuity", *Psychological Science*, Vol. 21 No. 5, pp. 661-666.

McCullough, M.E., Bellah, C.G., Kilpatrick, S.D. and Johnson, J.L. (2001), "Vengefulness: relationships with forgiveness, rumination, well-being, and the big five", *Personality and Social Psychology Bulletin*, Vol. 27 No. 5, pp. 601-610.

McCullough, M.E., Emmons, R.A., Kilpatrick, S.D. and Larson, D.B. (2001), "Is gratitude a moral affect?", *Psychological Bulletin*, Vol. 127 No. 2, pp. 249-266.

McCullough, M.E., Emmons, R.A. and Tsang, J.-A. (2002), "The grateful disposition: a conceptual and empirical topography", *Journal of Personality and Social Psychology*, Vol. 82 No. 1, pp. 112-127.

Nichols, H. (2016), "Caffeine: all you need to know", *from Medical News Today*, (accessed 24 April 2017)

Reivich, K. (2009), "Fishful thinking: cultivating gratitude in youth", *Communique*, Vol. 38 No. 1, pp. 1-2.

Salces-Cubero, I., Ramírez, E. and Ortega, A. (2018), "Strengths in older adults: differential effect of savoring, gratitude and optimism on well-being", *Aging & Mental Health*, Vol. 23 No. 8, pp. 1-8, doi: 10.1080/13607863.2018.1471585.

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, pp. 779-803, doi: 10.1108/JOCM-11-2019-0327.

Seligman, M.E.P. (2002), *Authentic Happiness: Using The New Positive Psychology To Realize Your Potential For Lasting Fulfillment*, Free Press, New York, NY.

Sergeant, S. and Mongrain, M. (2011), "Are positive psychology exercises helpful for people with depressive personality styles?", *The Journal of Positive Psychology*, Vol. 6 No. 4, pp. 260-272.

Shapiro, S.L., Schwartz, G.E.R. and Santerre, C. (2005), "Meditation and positive psychology", in Snyder, C.R. and Lopez S.J. (Eds), *Handbook of Positive Psychology*, Oxford University Press, Cary, NC.

Sheldon, K.M., Boehm, J.K. and Lyubomirsky, S. (2012), "Variety is the spice of happiness: the hedonic adaptation prevention (HAP) model", in Boniwell, I. and David, S. (Eds), *Oxford Handbook of Happiness*, Oxford University Press, Oxford, pp. 901-914.

Taylor, A.F., Kuo, F.E. and Sullivan, W.C. (2002), "Views of nature and self-discipline: evidence from inner city children", *Journal of Environmental Psychology*, Vol. 22 Nos 1/2, pp. 49-63.

Waters, L. (2012), "Predicting job satisfaction: contributions of individual gratitude and institutionalized gratitude", *Scientific Research*, Vol. 3 No. 12A, pp. 1174-1176.

Watkins, P.C. (2004), "Gratitude and subjective well-being", in Emmons, R.A. and McCullough M.E. (Eds), *The Psychology of Gratitude*, Oxford University Press, New York, NY, pp. 167-192.

Watkins, P.C., Uhder, J. and Pichinevskiy, S. (2014), "Grateful recounting enhances subjective well-being: the importance of grateful processing", *The Journal of Positive Psychology*, Vol. 10 No. 2, pp. 91-98, doi: 10.1080/17439760.2014.927909.

Watkins, P.C., Woodward, K., Stone, T. and Kolts, R.L. (2003), "Gratitude and happiness: development of a measure of gratitude and relationships with subjective well-being", *Social Behavior and Personality: An International Journal*, Vol. 31 No. 5, pp. 431-452.

Weinstein, N., Przybylski, A.K. and Ryan, R.M. (2009), "Can nature make us more caring? Effects of immersion in nature on intrinsic aspirations and generosity", *Personality and Social Psychology Bulletin*, Vol. 35 No. 10, pp. 243-259.

Wood, A.M., Froh, J.J. and Geraghty, A.W. (2010), "Gratitude and well being: a review and theoretical integration", *Clinical Psychology Review*, Vol. 30 No. 7, pp. 890-905, doi: 10.1016/j.cpr.2010.03.005.

Wood, A.M., Maltby, J. and Stewart, N.J.S. (2008), "Conceptualizing gratitude and appreciation as a unitary personality trait", *Personality and Individual Differences*, Vol. 44 No. 3, pp. 621-632.

Yang, K., Yan, W., Jia, N., Wang, Q. and Kong, F. (2020), "Longitudinal relationship between trait gratitude and subjective well-being in adolescents: evidence from the bi-factor model", *The Journal of Positive Psychology*, pp. 1-9, doi: 10.1080/17439760.2020.1818812.

Further reading

Emmons, R.A. and Crumpler, C.A. (2000), "Gratitude as a human strength: appraising the evidence", *Journal of Social and Clinical Psychology*, Vol. 19 No. 1, pp. 56-69.

Grant, A.M. and Gino, F. (2010), "A little thanks goes a long way: explaining why gratitude expressions motivate pro-social behaviour", *Journal of Personality and Social Psychology*, Vol. 98 No. 6, pp. 946-955.

Lambert, N.M., Fincham, F.D., Stillman, T.F. and Dean, L.R. (2009), "More gratitude, less materialism: the mediating role of life satisfaction", *The Journal of Positive Psychology*, Vol. 4 No. 1, pp. 32-42.

Layous, K., Sweeny, K., Armenta, C., Na, S., Choi, I. and Lyubomirsky, S. (2017), "The proximal experience of gratitude", *Plos One*, Vol. 12 No. 7, pp. 1-26.

Lin, C.C. (2013), "A higher-order gratitude uniquely predicts subjective well-being: incremental validity above the personality and a single gratitude", *Social Indicators Research*, Vol. 119 No. 2, pp. 909-924.

McCullough, M.E., Tsang, J.A. and Emmons, R.A. (2004), "Gratitude in intermediate affective terrain: links of grateful moods to individual differences and daily emotional experience", *Journal of Personality and Social Psychology*, Vol. 86 No. 2, pp. 295-309.

Podsakoff, P.M., MacKenzie, S.B., Paine, J.B. and Bachrach, D.G. (2000), "Organizational citizenship behaviors: a critical review of the theoretical and empirical literature and suggestions for future research", *Journal of Management*, Vol. 26 No. 3, pp. 513-563.

Podsakoff, N.P., Podsakoff, P.M., Mackenzie, S.B., Maynes, T.D. and Spoelma, T.M. (2014), "Consequences of unit-level organizational citizenship behaviors: a review and recommendations for future research", *Journal of Organizational Behavior*, Vol. 35 No. 1, pp. 87-119.

Podsakoff, N.P., Whiting, S.W., Podsakoff, P.M. and Blume, B.D. (2009), "Individual- and organizational-level consequences of organizational citizenship behaviors: a meta-analysis", *Journal of Applied Psychology*, Vol. 94 No. 1, pp. 122-141.

Solom, R., Watkins, P.C., McCurrach, D. and Scheibe, D. (2016), "Thieves of thankfulness: traits that inhibit gratitude", *The Journal of Positive Psychology*, Vol. 12 No. 2, pp. 120-129.

Spence, J.R. (2014), "Helpful today but not tomorrow? Feeling grateful as a predictor of daily organizational citizenship behaviours", *Personnel Psychology*, Vol. 67, pp. 705-738.

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

Computers and Electrical Engineering xxx (xxxx) xxx

Contents lists available at ScienceDirect

Computers and Electrical Engineering

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

De-CAPTCHA: A novel DFS based approach to solve CAPTCHA schemes☆

Aditya Atri [a], Ankita Bansal [b], Manju Khari [c], S. Vimal [d,*]

[a] Department of Information Technology, Netaji Subhash Institute of Technology, Sector 3, Dwarka, 110078, New Delhi, India
[b] Department of Information Technology, Netaji Subhash University of Technology, Sector 3, Dwarka, 110078, New Delhi, India
[c] Computer & System Sciences, JNU, New Delhi, India
[d] Department of Artificial Intelligence and Data Science, Ramco Institute of Technology, Rajapalayam, Tamilnadu, India

ARTICLE INFO

Keywords:
CAPTCHA
Island Problem
Depth-first Search
Security
Convolutional Neural Network

ABSTRACT

CAPTCHA stands for Completely Automated Public Turing Test to Tell Computers and Humans Apart. CAPTCHAs are used as security mechanism in web applications to differentiate between real users and automated users, also known as bots. Text-based CAPTCHAs are the popularly used CAPTCHA schemes due to their simplicity and thus, they are still being used despite the proposal of several attack mechanisms. In this work, the authors have proposed a novel approach to solve CAPTCHA schemes. In this approach, the authors have used Depth First Search algorithm for the extraction of characters from CAPTCHAs and Convolutional Neural Network for recognizing these extracted characters. The proposed approach was validated on 3000+ CAPTCHA schemes and proved to be efficient by providing an average accuracy of more than 92.0% in detecting CAPTCHA schemes.

1. Introduction

Internet has revolutionized everyone's life and in this inter-connected world information is like oil. With the advancements in Information and Communication Technology, huge amount of data is being generated and managed in digital form. Since there are many procedures and efforts to hack the system for sensitive data, therefore system security has emerged as a critical aspect of any web application [1,2]. Information is generally protected with various security mechanisms to prevent hacking approaches by automated systems called bots. CAPTCHA schemes are the most commonly used of these security mechanisms. A CAPTCHA scheme is typically an image containing different characters that one has to type in before entering or logging in the system or web application. Text based CAPTCHA schemes are used by websites of various popular Indian government institutions (like Aadhar-UIDAI, Digital India, IRCTC, SBI Online, Delhi Police, etc.) as well as the largest internet companies (like Amazon, Yahoo!, Microsoft, eBay, etc.). The primary purpose of CAPTCHA is to safeguard user logins against brute force password guessing by bots, to prevent automated usage of various services. But with the introduction of deep learning and computer vision, CAPTCHAs can now often be defeated by bots [3,4]. While CAPTCHAs have evolved through more complex images to address these risks, wide deployment and use of text-based CAPTCHAs and their increasing vulnerability can result in serious consequences. That is why, through this work, authors have proposed a novel approach to solve complex CAPTCHA schemes. The proposed approach attempts to highlight risks with even the more complex CAPTCHA schemes and raises questions on how to re-design the CAPTCHAs to enhance security.

☆ This paper is for special section VSI-mlsec. Reviews were processed by Guest Editor Dr. Suyel Namasudra and recommended for publication.

* Corresponding author.
E-mail addresses: adityaa.it.16@nsit.net.in (A. Atri), ankita.bansal06@gmail.com (A. Bansal), manjukhari@yahoo.co.in (M. Khari), svimalphd@gmail.com (S. Vimal).

https://doi.org/10.1016/j.compeleceng.2021.107593
Received 3 January 2021; Received in revised form 10 June 2021; Accepted 4 November 2021
0045-7906/© 2021 Elsevier Ltd. All rights reserved.

Automated breaking of CAPTCHAs is not new and several researchers and scientists have proposed different methods to break text-based CAPTCHAs. The most important thing in CAPTCHA breaking is the extraction of characters from the CAPTCHA schemes [5]. Existing algorithms to break CAPTCHAs are typically hard-coded and work for some specific CAPTCHA schemes like Dynamic Cognitive Game (DCG). However, the main limitation of the existing algorithms is that they are neither generic nor robust in detecting slight changes in the CAPTCHA schemes. In addition to this, one key constraint with existing algorithms is the requirement of large training datasets and a labor-intensive approach [6]. Existing algorithms fail with a slight change in CAPTCHA configuration and they break CAPTCHAs in non-linear time with less accuracy. Works in literature claimed that their generic methods can solve a large family of text-based CAPTCHAs with ever-changing CAPTCHA schemes [7–11]. However, CAPTCHAs have evolved further and various resistance mechanisms, such as noise arcs, complicated backgrounds and two-layer structures have been added to make them difficult to be attacked by early generic methods. Moreover, increasingly websites have evolved to send random CAPTCHAs with different design features to users. Therefore, it is clear that the creation of a generic and dynamic attack method is needed which can break diverse CAPTCHA schemes. In research work [7], combining segmentation and recognition decreased the time to break the CAPTCHAs significantly, however this was achieved at the cost of reduced accuracy. There is lack of proper balance between accuracy and speed. In research work [8], the Log-Gabor filters are used to extract character components from CAPTCHA images along four directions but the algorithm proposed in this work based on Depth First Search (DFS) focuses on all the directions i.e. 8. The work on Recurrent Neural Network (RNN) which includes less feature compatibility than Convolutional Neural Network (CNN), is not suitable for spatial data such as in CAPTCHA images is discussed in [9]. Thus, the main motivation behind this work is to promote research and development in security mechanisms arena to ensure more robust authentication features. The focus is to protect the user's data and privacy using improved mechanism while maintaining the simplicity. We have tried to break the core idea behind CAPTCHAs and aim to develop new and better authentication mechanisms like reCAPTCHAs. As CAPTCHA's verification is not a rule-based solution, this inspires us to use machine learning. Machine Learning helps in finding patterns in the chaos of unstructured big datasets like CAPTCHA images. This paper proposes a simple, dynamic, and effective text-based CAPTCHA solver based on DFS for the extraction of characters from CAPTCHA schemes and CNN for recognizing them, hence providing a reliable approach to solve CAPTCHA schemes with greater speed and accuracy. The major contribution of this research paper is as follows :

1. Removal of noise from the background and extraction of characters from CAPTCHA schemes.
2. Analysis of the performance of the proposed CAPTCHA solver when different security features like overlapping, rotation, distortion and waving are present in CAPTCHA schemes.
3. Comparison of proposed CAPTCHA solver (DFS+CNN) with the approaches proposed in the literature.

To the best of the authors' knowledge, this approach has departed from all the shortcomings present in previous work [7–11]. Before recognition, we have removed the noise from the background using the DFS approach in the extraction part that makes the process easier and effective. We have thoroughly discussed the approach of the extraction of characters from the CAPTCHA schemes using the DFS algorithm. We have included all kinds of security features like overlapping, rotation, distortion, and waving present in the CAPTCHA schemes to make the proposed approach more robust and generic. We have collected 1000+ CAPTCHA schemes from 30+ websites which include IRCTC, UIDAI, SBI Online, AliBaba, etc. In this way, we validated the proposed CAPTCHA solver on more than 30,000 CAPTCHA schemes. The proposed approach is found to be highly efficient as it solves most of the CAPTCHA schemes within 50.00 ms using a personal computer (local GPU). The proposed algorithm present in this paper handled background noises and complexity very efficiently. We have achieved an average accuracy rate of 98.0%, 90.6%, and 71.9% for easy, medium, and complex CAPTCHA schemes. Our proposed approach raises questions for the popular websites (IRCTC, UIDAI, SBI Online, AliBaba) which are employing the use of CAPTCHAs on how to redesign the existing CAPTCHAs for ensuring security in future. We have provided a link to demo: https://www.youtube.com/watch?v=69tpmyV2lzs and implementation: https://github.com/AdityaAtri/captcha-breaker.

The organization of this paper is as follows: In Section 2, we investigated various studies which have proposed different techniques to break CAPTCHAs. We have discussed different features associated with the text-based CAPTCHA schemes used by various Indian Government websites and MNCs in Section 3. In Section 4, the extraction approach based on DFS and model based on CNN to recognize the extracted characters from CAPTCHA schemes is discussed. In Section 5, we validated the proposed approach on various CAPTCHA schemes of popular websites. Later in the section, we compared the results of state-of-art-attacks with the discussed approach. Next section presents conflict of interest. Finally, we conclude this paper with limitations and future works in Section 6.

2. Related work

In this section, we discuss some of the popular works on text-based CAPTCHAs. Text-based CAPTCHAs are most popular and are widely used by several websites because of their easiness [12]. Text-based CAPTCHAs are undergoing a rigorous process of improvement similar to watermarking and cryptography. Due to the increase in cyberattacks, it became necessary to propose more robust and strong CAPTCHA schemes. Although, there is literature and work has been done to improve the power of CAPTCHA solvers to handle text based CAPTCHA schemes, but the CAPTCHA solvers have shown varied performance. Thus, at present a conclusion cannot be drawn in favor of a particular CAPTCHA solver and more research is required in this field. This has motivated the authors to propose a CAPTCHA solver to handle text-based CAPTCHA schemes. The work by Mori et al. was among the first attempts to break the text-based CAPTCHA schemes [13]. Breaking of most of the CAPTCHAs from Captchaservices.org was done by Yan et al. with the help of a simple character segmentation method which was based on the counting of the number of pixels of individual characters [14]. Further, they developed an improved method of segmentation that performed an attack on early CAPTCHA schemes used by Yahoo!, Microsoft and Google [15]. Another work presented by Gao et al. targeted CAPTCHA schemes

Table 1
Brief summary of various related studies.

Study in literature	Work done	Advantages	Disadvantages
Alaa Thobhani et al. [11]	Proposed CNN model which works on binary images to solve CAPTCHAs. The model is based on making the number of copies of input CAPTCHA image equal to the number of characters in that CAPTCHA image.	Focuses on low complexity and avoiding segmentation step using attached binary images algorithm	Extreme resource intensive and overhead in creating number of copies equal to number of characters present.
Yang Zhang et al. [23]	Surveyed different techniques for breaking/solving of text-based CAPTCHAs.	Thoroughly discussed the basic principles, advantages, and disadvantages of the existing methods	Not enough transparency with results and datasets
Guixin Ye et al. [17]	Trained a base solver using synthetic CAPTCHAs generated using CAPTCHA synthesizer and then used transfer learning to fine-tune the base solver to work on real CAPTCHAs.	Fewer real and manually-labeled CAPTCHAs and improved performance	The studies in literature have shown that adversarial examples produced by GANs can confuse machine learning classifiers easily.
Kat Krol et al. [12]	Gathered users' experience for 3 human verification mechanisms: reCAPTCHA, PlayThru and NoBot. Also, suggested a Possible Replacement Technology.	Focused on both qualitative and quantitative aspects	Participants could be biased while engaging with realistic primary tasks
Shotaro Sano et al. [18]	Proposed a framework for solving continuous CAPTCHAs and measured its efficiency using actual CAPTCHA questions from Google's reCAPTCHA.	Guidelines to prevent possible attacking from HMM-based CAPTCHA solvers and revealing vulnerabilities in continuous CAPTCHAs	User testing on each noise and/or distortion technique is not discussed

of hollow characters [16]. A Generative Adversarial Network-Based Approach to solve the CAPTCHA is presented by Guixin Ye et al. [17]. HMM-based CAPTCHA Breaker for Overlapped Symbols is presented by Shotaro Sano et al. [18]. Early methods to break Google reCAPTCHA is shown by Suphannee Sivakorn et al. [19]. Another work develops a simple but generic attack for Text-Based CAPTCHAs employed by machine-learning based classifier [10]. As a final remark, it is important to discuss the work based on the foundations of the CNN machine learning approach. This technique becomes useful in developing applications to escape from spam mail filtering [20], bypass malware detection [21], and confuse the machine learning-based classifiers [22]. A survey of research on CAPTCHA designing and breaking techniques was conducted by Yang Zhang et al. in 2019 showing the scientific and technological progress in both the design and attack of CAPTCHAs [23]. A Deep-learning-based CAPTCHA breaking system with low complexity and avoiding segmentation step using attached binary images algorithm is proposed by Alaa Thobhani et al. in 2020 [11]. The brief summary of all the related studies is shown by Table 1.

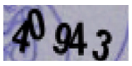
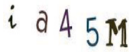
To the best of authors' knowledge, no work has depicted the DFS technique to extract characters from the CAPTCHA schemes and construct a generic and dynamic CAPTCHA solver for text-based CAPTCHAs. The DFS technique to extract the characters from the CAPTCHA schemes makes the CAPTCHAs solver more effective. The time complexity of DFS algorithm to process the attached nodes is linear time which is much lesser than exponential complexity discussed in [10]. The DFS algorithm only processes the neighbor nodes matching the required conditions. Hence, it converges the search area and neglects the noisier(unmatched) nodes leading to higher accuracy. It focuses on all the neighboring pixels of CAPTCHA schemes i.e. 8 unlike Log-Gabor filters which focuses only in 4 directions as presented in [8]. Thus, the discussed approach of DFS helps in processing the input CAPTCHA in less time and with higher accuracy. The proposed approach has been proved effective, efficient, dynamic, and simple as compared to the already proposed techniques.

3. Research background

This section explains various features associated with CAPTCHA schemes. Features are used to make the CAPTCHA schemes more tolerant to cyberattacks. Following are some of the important features which are included in most of the widely used CAPTCHA schemes:

1. CAPTCHA must contain (lowercase and uppercase) alphabets and numeric characters. Multiple randomizing functions (JavaScript) are used to generate random CAPTCHA to make it not susceptible to attack.
2. The length of CAPTCHA is varied from 4 to 6.
3. The text inside CAPTCHA is rotated at various angles.
4. Multiple font types are included to prevent the attack.
5. Lines are used to preventing segmentation.
6. The numbers and the length of lines and their positions are varied each time to distort the text image plus including the waving nature of characters.
7. The text image is blurred using a specific technique to make CAPTCHA difficult to break.
8. Image dimensions are varied inconsistently and line's colors are kept in grayscale colors but at a different level for each time.

Table 2
Various features associated with Text-Based CAPTCHA schemes.

Website	CAPTCHA image	Security features	Complexity
UIDAI		Alphabetic and numeric characters, colored background, varied font sizes, waving	Medium
IRCTC		Numeric characters, overlapping characters, waving	Easy
SBI Online		Alphabetic and numeric characters, overlapping characters, distorted characters, varied font sizes, waving, occluding lines	Complex
AliBaba		Alphabetic and numeric characters, overlapping characters, waving, distorted characters, varied font sizes, waving	Complex
Digital India		Alphabetic and numeric characters, varied font styles, varied font sizes, rotation, waving	Medium
Delhi Police		Alphabetic and numeric characters colored dotted background, varied font sizes, waving	Medium
India.gov.in		Alphabetic characters, overlapping character, rotating, distortion, varied font sizes, waving	Complex
CAT 2019		Alphabetic characters, varied font size, rotation, colored background, distorted characters	Complex

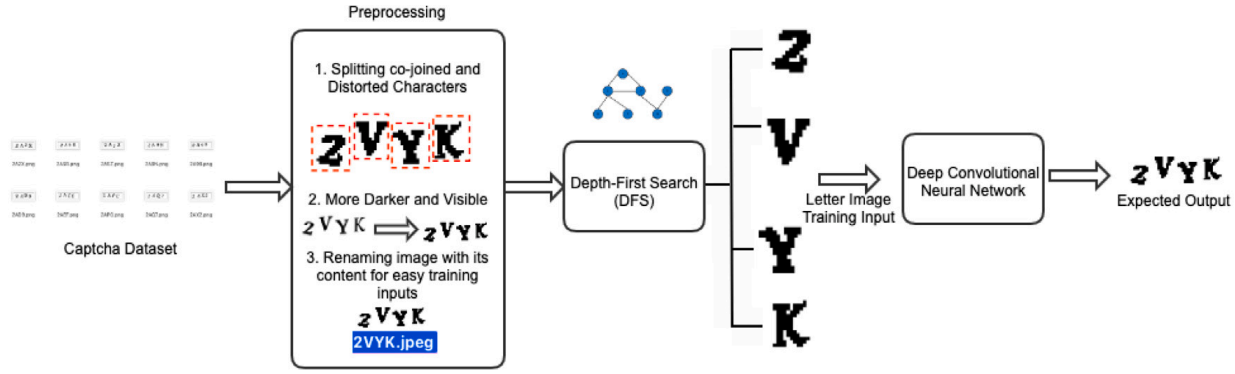


Fig. 1. The proposed algorithm's flowchart.

Since, these are the important features which are present in the text-based CAPTCHAs, the authors in this paper designed a CAPTCHA solver which is able to solve a CAPTCHA having these above stated features. For the purpose of validation, a large number of CAPTCHAs from various Indian government and MNCs websites are taken. For the purpose of demonstration, a CAPTCHA image is extracted from these websites and the security features associated with each image is shown in Table 2. In addition to this, each CAPTCHA image is categorized as either simple, medium, or complex. CAPTCHAs are considered simple when there is no presence of overlapping, rotation, distortion and waving nature. The medium CAPTCHAs contains Overlapping and Waving nature. The complex CAPTCHAs contain all features like overlapping, rotation, distortion, waving along with occluding lines. This categorization can be clearly seen in Table 2. The novel technique proposed in this paper is able to break even the complex type of CAPTCHA with good accuracy.

4. The proposed scheme

This paper proposes a novel approach to develop an efficient CAPTCHAs solver. The proposed algorithm follows broadly two main steps: (1) extraction of characters using the DFS algorithm and (2) recognition of extracted characters using CNN. Each of these steps is explained in detail in this section supported by suitable figures and examples. The flowchart of the proposed algorithm is shown by Fig. 1. This proposed scheme is implemented on Django Framework where machine learning algorithms to break CAPTCHA schemes run in the backend. Different technologies like Selenium to scrape various schemes from the web are used.

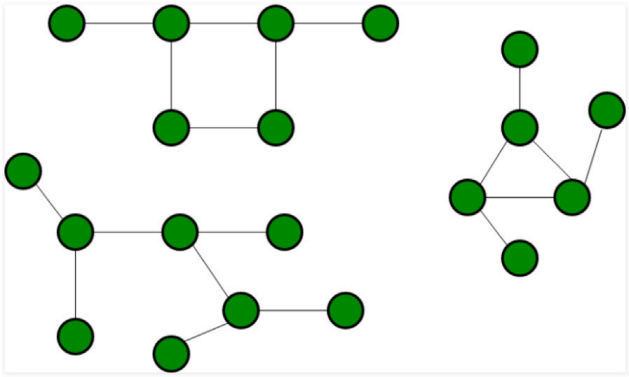


Fig. 2. Strongly Connected Components of a Graph.

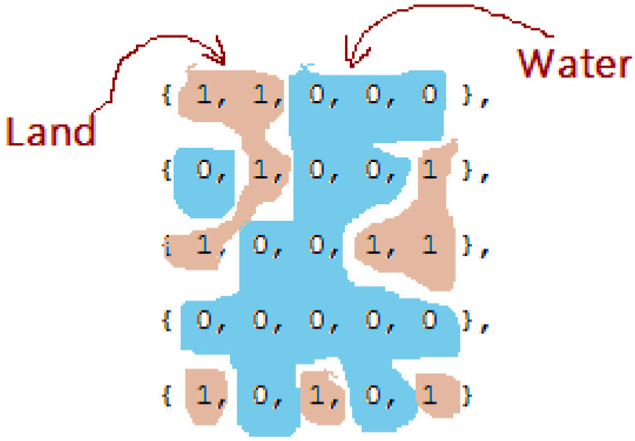


Fig. 3. Identification of Islands.

4.1. Extraction of characters using the depth-first search (DFS) algorithm

In Computer Science, DFS algorithm is a recursive algorithm that focuses on an exploration of a neighboring node before the source node itself. It involves exhaustive searching of all the nodes by going ahead and coming to the same node by backtracking.

Depth First Search Algorithm :

```
DFS-recursive ( G, s ) :
   visited [s] = True
   for ( i = 0  to size( adjacency_list[u]) :
          v = adjacency_list [u][i]
          if ( visited [v] == False ) :
                 DFS ( G, v )
```

Here "G" is a graph and "s" is a source vertex. The vertex "s" is marked True if visited. Then, it explores the immediate connected nodes and call the DFS function with source node as 'v" which is not visited yet. Before the introduction of extraction-algorithm, there is need to understand the very famous problem in graphs : The Island Problem — given a boolean matrix.

On an island, we can move to any point from any point but cannot reach another island. It means all the points on an island are strongly connected. Fig. 2 shows the connected components of an undirected graph. We start with one node and traverse all those nodes which are strongly connected. In this way, we can count the number of islands present in the sea. So, the way in which we identified the islands as shown in Fig. 3, we will identify the characters from the CAPTCHA schemes. The input and output of DFS algorithm can be defined as :

Input of DFS Algorithm: Darker and visible CAPTCHAs

Output of DFS Algorithm: Separated CAPTCHAs letters

Next, this paper will discuss different scenarios when CAPTCHAs do not have overlapping characters and when CAPTCHAs have overlapping characters.

-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	-1	-1	1	-1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1

Fig. 4. Pixel matrix of character X.

Fig. 5. CAPTCHA scheme having co-joined characters.

Case 1: CAPTCHAs having no overlapping of characters

Consider a simple CAPTCHA scheme in which a single character is composed of the same pixel values. Although, different characters may have different pixel values to make the CAPTCHA schemes more robust. Suppose there is character 'X' in which all pixels have the same value (say 1) and surrounded with background pixel value, (say -1) as shown in Fig. 4. Both character's pixel value and background's pixel value cannot be the same otherwise we cannot see the character in the CAPTCHA schemes. Now, start from the leftmost pixel of character 'X' and traverse all those pixels which are strongly connected having the same pixel value of 1. If there is no overlapping, then we can easily traverse all the pixels present in the character 'X' without going to the next character. If the next character's pixel value is not the same as that of the previous character's pixel value, then we do not even need to consider the overlapping case because the DFS approach will only traverse the pixels having the same values. If the next character's pixel value is the same as that of the previous one, then the DFS algorithm can go to the next character and we need to consider the overlapping case as different character has started.

Case 2: CAPTCHAs having overlapping of characters

If there is overlapping and the next character has a different pixel value than the previous one, then we can simply apply the algorithm discussed in Case 1. But when there is overlapping in the CAPTCHA as shown in Fig. 5 and the next character has the same pixel value as that of the previous one, then it is very important to separate the two characters. If we do not handle this case, we will end up in the wrong predictions that will decrease the accuracy. If a single contour area that we covered with DFS algorithm is much wider than its height, it implies that the letters are overlapping. To solve this, we split the conjoined letter in half down the middle and consider it as two different letters as shown in Fig. 6. Joining more than two letters in a CAPTCHA scheme is very much rare as it will hamper the user experience. Once we are able to extract the characters from the CAPTCHA, the training of the model to recognize the character is simple. We can also decrease the pixel value to make the character more visible and darker as shown in Figs. 7 and 8. This made extraction procedure easy.

4.2. Recognition of extracted characters using convolutional neural network

This paper adopted the CNN Multi-Layered Perceptron (MLP) as CNN works best for spatial data such as images. CNN is build on feed-forward artificial neural network [24] with variations of multi-layer perceptrons designed to require less amount of pre-processing. Krizhevsky et al. [25] proved that the usage of CNNs gives higher image classification accuracy when compared to other

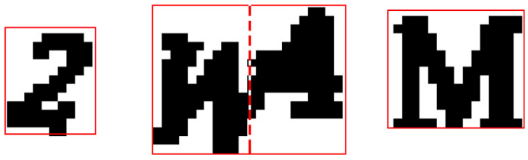


Fig. 6. CAPTCHA scheme having co-joined character with Split.

Fig. 7. CAPTCHA schemes with less dark pixels.

Fig. 8. CAPTCHA schemes with more dark pixels.

Fig. 9. Training dataset required by the proposed approach.

machine learning techniques such as Support Vector Machines. We have not used fully connected layers because weights required by fully connected layers are huge and hence, it hampers processing and results in less accuracy. In CNNs, the neurons in layers are only connected to a small region of layers before them. Hence, we need to handle less amount of weights. Now, we will move from creating dataset to recognition of characters in Captcha schemes:

4.2.1. Creating the datasets

To train any machine learning system, we need training dataset. We created a web scrapper using Python Selenium library to download different kinds of CAPTCHAs making dataset more inclusive and versatile. To deal with noise, we converted various Portable Network Graphics (PNGs) into Joint Photographic Experts Groups (JPEGs) as JPEGs are of small size. We manually removed those CAPTCHAs from dataset if they have different languages other than English numeric and alphabets. We removed CAPTCHAs having highly distorted characters which are not humanly understandable or produce biasness in prediction. We used python Computer Vision (CV) cv2.copyMakeBorder() method to create a border around the image like a photo frame. We resized the every CAPTCHA image to uniform size of $(20 * 20 * 1)$ that will later used by neural model. To break a CAPTCHA scheme, we want training data as shown in Fig. 9. It can be seen from Fig. 9 that we have manually created a training dataset with 5000 PNG files with filename same as the characters inside it. For the purpose of replication, this dataset is uploaded on Github and can be accessed from the link : https://tinyurl.com/2rx6a7we. In other words, we named the CAPTCHA images with the same text that is present inside the images. It would help in the training of the model and predicting the results with greater accuracy.

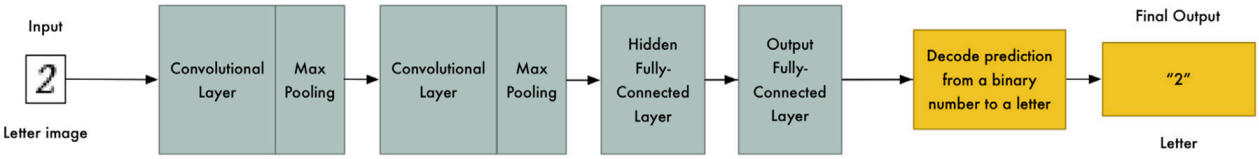


Fig. 10. Prediction of the text inside the CAPTCHA schemes using a DCNN having input as a single letter.

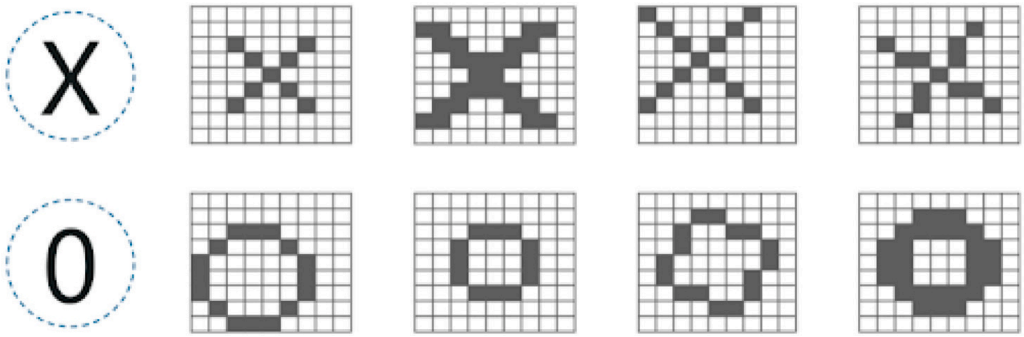


Fig. 11. Different versions of characters that must be recognized by CNN model.

4.2.2. Simplifying the problem

Now we have obtained the dataset as we wanted. We could have used these images to directly train the model using Deep Convolutional Neural Network (DCNN). However, we try to make the problem much simpler which will lead to saving of resources (less training data and computational power will be required).

We have converted each CAPTCHA image to grayscale for easier processing. Then we have added padding to avoid any mismatch in predication of characters. Next, we split the image such that each letter is a separate image on which neural network can be trained. In other words, we have to train neural network to recognize only a single letter at a time. This extraction of characters from the CAPTCHA schemes is already discussed in the previous subSection 4 (A). For extraction, we also took help of Contours function in Python OpenCV library. All the extracted letters from CAPTCHA schemes are sorted based on the their presence on x-coordinate to make sure we are processing them left-to-right to match the right character's image with filename. So, the extraction part supplemented the recognition part. Now we only need to recognize images of single letters and numbers of English Alphabet and hence, we do not need any complex neural network architecture. Since, we are recognizing images of single letter, we will use a simple CNN architecture with two convolutional layers and two fully-connected layers as shown in Fig. 10.

4.2.3. Building and training the model

After applying the DFS algorithm to extract characters from CAPTCHA schemes, we have reduced the problem from recognizing the images to recognizing letters which is a much easier problem to solve. Now, we only need to recognize images of single letters and numbers, hence, we do not need a very complex neural network architecture.

CNN has the following layers

1. Convolutional
2. Rectified Linear Unit (ReLu) layer
3. Pooling
4. Fully Connected

We have already discussed that CNN is best for image classification and is best for digit-classification. The trained model is used to identify the extracted letters. There may be different versions of characters that must be recognized by the model as shown in Fig. 11. As we may have distorted, rotated as well as waved characters in the CAPTCHA schemes, we have to deal with different versions of the character to make this model more robust and effective. We may also get a lot of missing values if we compare a normal image with another 'x' rendition.

How to fix this problem?

We consider a small part of the pixels called filters. We try to match them in the corresponding nearby locations to see if we get a match as shown in Fig. 12. By doing this, CNN gets a lot better at seeing similarities than directly trying to match entire image. Now, we will discuss the functionalities of each layer of CNN during character recognition:

(a) Convolutional Layer. In this layer, we perform the operation of "convolution" where we move the features to all the possible positions on the image's pixel matrix. Three filters used for detecting alphabet say 'X' are shown in Fig. 13. We explain the steps used to perform "convolution".

First, we multiply every pixel of the feature matrix with every corresponding pixel present in the existing matrix and the result is stored in the buffer matrix as shown in Fig. 14. Next, we add all the values present in the buffer matrix. This sum is divided by

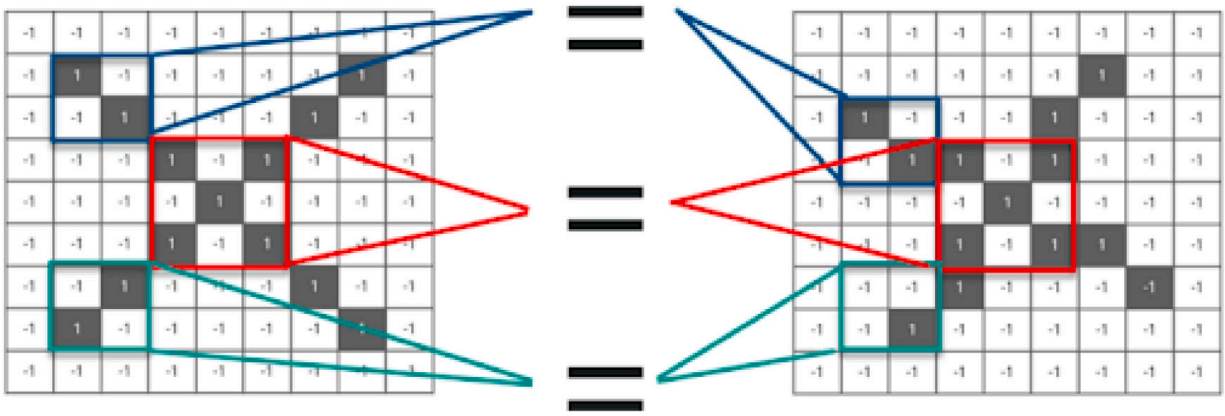


Fig. 12. Comparing the parts of the input character with standard character instead of comparing the character as a whole.

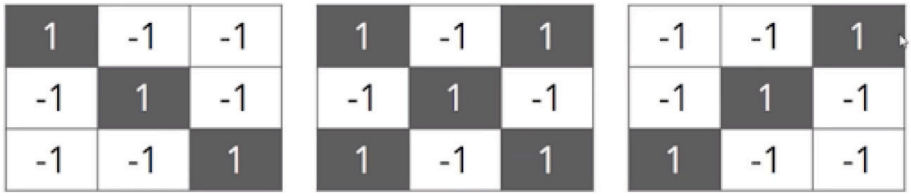


Fig. 13. Different types of the filters used for detecting the standard alphabet 'X'.

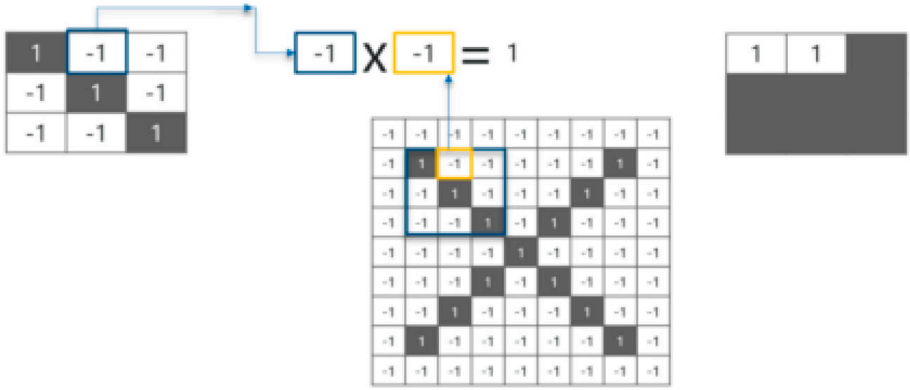


Fig. 14. Storing of result into buffer matrix.

the total number of pixels in the feature image and we keep this value at the center of the filtered matrix. Now again, move this filter and perform the same operation as mentioned above at any pixel in the image. We see how the feature matches the area of the image. We get the output as filtered matrix as shown in Fig. 15. Here we considered just one filter.

(b) ReLu Layer. ReLu layer works according to the function defined in Eq. (1).

$$ReLuLayer(x) = \max(0, x) \tag{1}$$

The parameter x is defined as any two dimensional x-axis point. According to this equation, if input is negative, the output becomes 0, else input has a linear relationship with output. Thus, using this equation, all the negative values from the filter matrix are replaced with zero as shown in Fig. 16. This is done so that the positive and negative values do not get added up to give sum as 0.

(c) Pooling Layer. In this layer, we shrink the image size into a smaller size.

Steps :

1. Take a window size (usually 3 or 4)
2. Take a stride(usually 2)
3. Move the window across the filtered matrix
4. From every window, take out the maximum value

0.77	-0.11	0.11	0.33	0.55	-0.11	0.33
-0.11	1.0	-0.11	0.33	-0.11	0.11	-0.11
0.11	-0.11	1.0	-0.33	0.11	-0.11	0.55
0.33	0.33	-0.33	0.55	-0.33	0.33	0.33
0.55	-0.11	0.11	-0.33	1.00	-0.11	0.11
-0.11	0.11	-0.11	0.33	-0.11	1.00	-0.11
0.33	-0.11	0.55	0.33	0.11	-0.11	0.77

0.77	0	0.11	0.33	0.55	0	0.33
0	1.00	0	0.33	0	0.11	0
0.11	0	1.00	0	0.11	0	0.55
0.33	0.33	0	0.55	0	0.33	0.33
0.55	0	0.11	0	1.00	0	0.11
0	0.11	0	0.33	0	1.00	0
0.33	0	0.55	0.33	0.11	0	1.77

Fig. 15. Output obtained after applying the ReLu function on one of the filtered matrix.

0.77	0	0.11	0.33	0.55	0	0.33
0	1.00	0	0.33	0	0.11	0
0.11	0	1.00	0	0.11	0	0.55
0.33	0.33	0	0.55	0	0.33	0.33
0.55	0	0.11	0	1.00	0	0.11
0	0.11	0	0.33	0	1.00	0
0.33	0	0.55	0.33	0.11	0	1.77

1.00	0.33	0.55	0.33
0.33	1.00	0.33	0.55
0.55	0.33	1.00	0.11
0.33	0.55	0.11	0.77

Fig. 16. Reduction of filtered matrix size after applying max-pooling.

It is important to note that we started out with a $7 * 7$ matrix but now the same matrix after pooling reduced to 4*4 as shown in Fig. 16. This is for only one filter. We have to perform it for the other two filters as well. In this way, we get 3 reduced matrices. But, we can further reduce these matrices. We apply the previous three operations to get $2 * 2$ matrices.

(d) Fully Connected Layer. The last layer used in the model is fully connected layer where actual classification takes place. Here, all the possible paths from input to output are considered because we have significantly reduced the size of the matrix from $7 * 7$ to $2 * 2$. This will reduce the demand for the total number of weights required in the neural network. Next, we take all such reduced matrices and insert them in a single vector known as a classification vector as shown in Fig. 17. In the classification vector, some values are 1 which are considered as high. For every character, there are different elements that are high. It can be seen from Fig. 17 that if the 1st, 4th, 5th, 10th, and 11th, values are high, we can classify the input as 'X'. This concept is extended to other alphabets and numbers. If values are arranged in some particular manner, they can be mapped to an actual alphabet or number which we require. We have provided the dataset, code and video demo (https://www.youtube.com/watch?v=69tpmyV2lzs and https://github.com/AdityaAtri/captcha-breaker).

5. Results evaluation and discussion

In this section, we have evaluated and validated the approach using 3000+ CAPTCHA schemes from 30+ websites. We have also compared the results with approaches present in literature. The results are explained in detail in this section supported by suitable tables and graphs. We developed a web application on which different CAPTCHA schemes were embedded to calculate the exact amount of time required by model to break it as presented by Fig. 18 (https://www.youtube.com/watch?v=69tpmyV2lzs). We made a login system where the computer automatically fills the CAPTCHA and enters into the system without human interference. This shows how CAPTCHA can be broken using the algorithm proposed by authors in this paper.

5.1. Validation of the proposed CAPTCHA solver

We have shown the Accuracy Rate (AR) of the developed approach (extraction + recognition) on various popular websites. We have taken one sample due to space constraint from every website (mainly 8) to present the expected output, observed output, AR and running time per CAPTCHA. The performance of the proposed solver is shown in Table 3 in terms of the AR and the running time (in ms) taken to solve each CAPTCHA. The AR shown in the Table 3 is the accuracy achieved when the number of CAPTCHA schemes reaches 1000. For each website, which include IRCTC, UIDAI, SBI Online, AliBaba, etc., we have collected 1000+

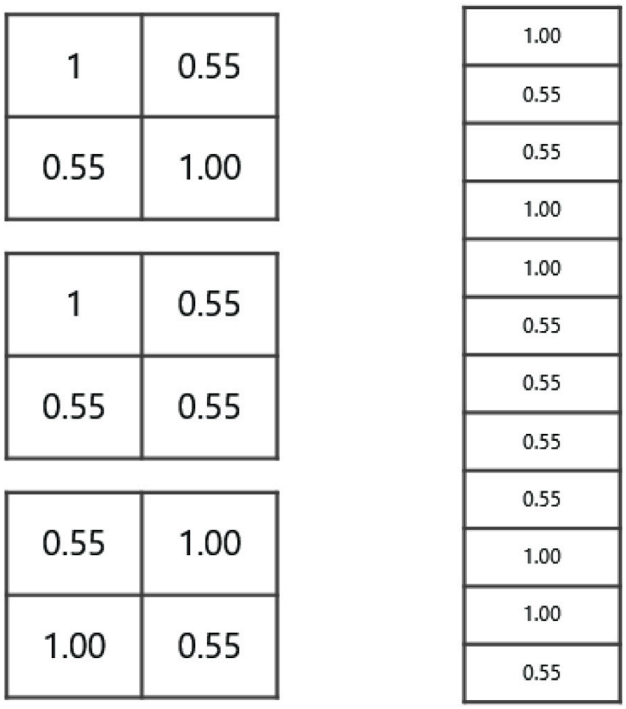


Fig. 17. Standard classification vector for character 'X'.

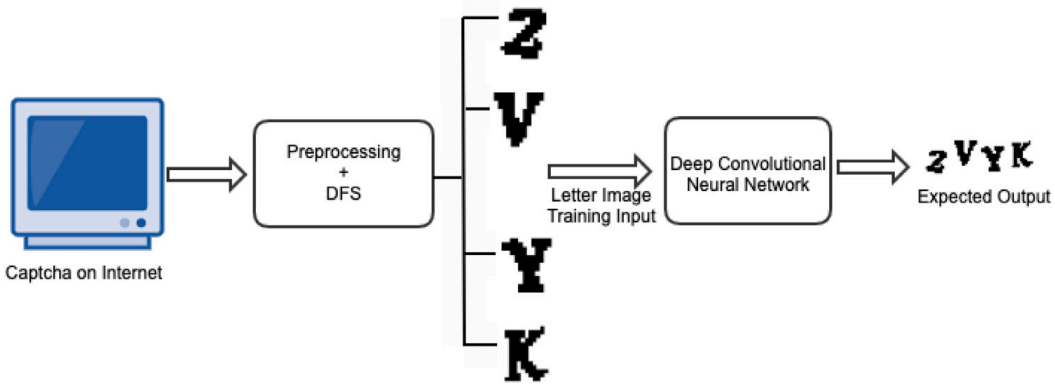


Fig. 18. Prediction of input CAPTCHA scheme using proposed method.

CAPTCHA schemes. In this way, we performed the approach on more than 30,000 CAPTCHA schemes. Below formula (2) shows how to calculate AR :

$$AR = \frac{\text{Correctly identified CAPTCHA Schemes} * 100}{\text{Total number of CAPTCHA Schemes for any particular website i.e. 1000}} \tag{2}$$

The running time (ms) per CAPTCHA is calculated as total time required to process the given input CAPTCHA image. Below formula (3) calculates the running time (ms) per CAPTCHA using time() function of Python library:

$$CaptchaRunningTime\ (ms) = \text{Output Time} - \text{Input Time} \tag{3}$$

Output time is defined as the time (ms) at which model decoded the CAPTCHA input. Input time is defined as the time (ms) at which CAPTCHA image is provided to model as an input. For UIDAI, the expected output is "5gzpf" and observed output is "5gzpf" which is solved with AR of 92.0%. For IRCTC, the expected output is "50623" and observed output is "50623" which is solved with AR of 98.9%. For the IRCTC, the CAPTCHA schemes are easy containing only numerical characters. SBI Online CAPTCHA schemes are solved with AR of 86.6% having expected output as "wcrg2" and observed output as "wcrg2". The AR is less in case of SBI and AliBaba CAPTCHA schemes as compared to simple (AR — 98.9%) and medium (AR — 92.0%) CAPTCHA schemes due to more complexity. Although, reasonable high AR as compared to [7–11] in complex CAPTCHA schemes shows efficiency of the CAPTCHA solver. Similarly, CAPTCHA schemes of CAT 2019 is solved with AR of 78.0% where expected output is "V7sbfve" and observed

Table 3
Performance of the proposed CAPTCHA solver.

Website	CAPTCHA image	Expected CAPTCHA output	Observed CAPTCHA output	Proposed approach AR (%)	Running time per CAPTCHA (ms)
UIDAI		5gzpf	5gzpf	92.0%	43.80
IRCTC		50623	50623	98.9%	30.20
SBI Online		wcrg2	wcrg2	86.6%	46.80
AliBaba		6EBT	OEBI	67.3%	41.57
Digital India		WQBe	WQBe	98.0%	40.63
Delhi Police		nVZhgA	nVZhgA	82.0%	45.40
India.gov.i		ZT6YL	KT6YL	56.0%	40.32
CAT 2019		V7sbfve	V7sbfve	78.0%	51.34

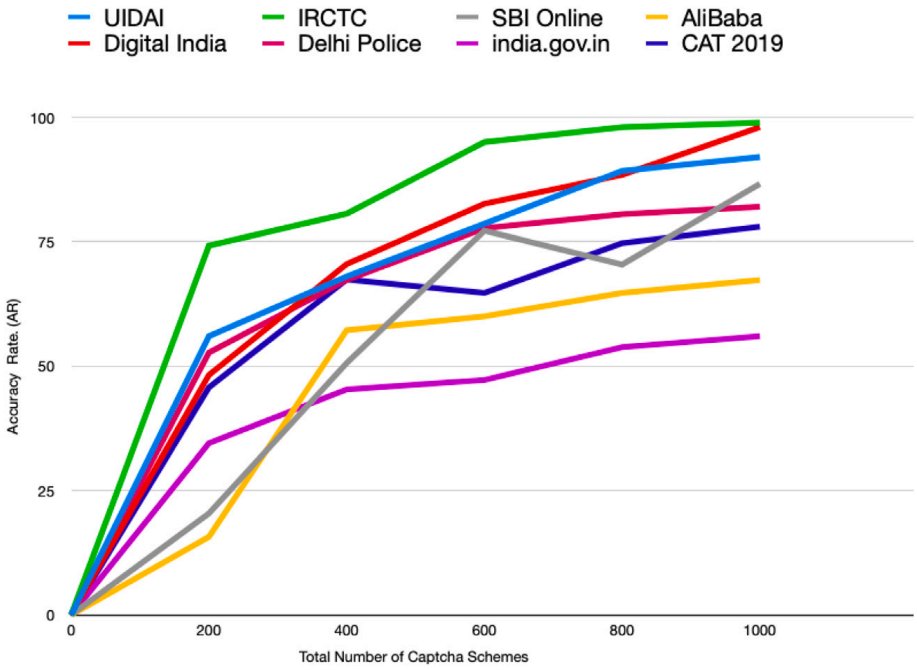


Fig. 19. Relationship between the number of CAPTCHAs and the accuracy rate.

output is "V7sbfve". On an average, this model solved the simple CAPTCHA schemes like IRCTC within 30.20 ms with average AR of 98.0%. The medium complexity CAPTCHA schemes like UIDAI, Digital India, Delhi Police etc. are solved between 40.00 to 45.00 ms with average AR of 90.6%. The difficult ones like AliBaba, SBI Online, CAT 2019, India.gov.in etc. are solved within 45.00 to 50.00 ms with average AR of 71.9%.

The relationship between the number of CAPTCHA schemes for any particular website and AR obtained is shown in Fig. 19. We measured the variation in performance of model by changing number of samples in training/testing dataset. In general, with the increase in CAPTCHA schemes, the AR increased.

For every CAPTCHA schemes except SBI Online and CAT 2019, the AR increased with increase in the total number of CAPTCHA schemes. There was little drop down in the AR of SBI between 600–800 CAPTCHA schemes possibly due to distorted CAPTCHAs and highly increased complexity. Similar is the case with CAT 2019 CAPTCHA schemes.

5.2. Impact of various security features

There are 4 common security features for text-based CAPTCHAs: overlapping, rotation, distortion, and waving. The different combinations of these security features in CAPTCHAs make them more complex and hard to break. The authors have included these security features to show that the CAPTCHA solver proposed in this paper is able to break CAPTCHAs with different security features. For the time being, we have excluded features like occluding lines and noisy background. Table 4 shows the AR given by the proposed model when different security features are present.

1. Overlapping : By decreasing the space between adjacent characters, the efficacy of solver decreases. If the gap between characters is reduced to 4, 6, 8, and 10 pixels, the AR of the developed solver is 68.0%, 50.4%, 46.0%, and 15.0% respectively.

Table 4
Ability of the proposed CAPTCHA solver to deal with different security features.

No.	CAPTCHA image	Overlapping	Rotation	Distortion	Waving	Proposed approach AR (%)
1.		Yes	Yes	No	No	78.5
2.		Yes	No	Yes	No	75.2
3.		Yes	No	No	Yes	86.5
4.		No	Yes	Yes	No	67.2
5.		No	Yes	No	Yes	87.5
6.		No	No	Yes	Yes	66.6
7.		Yes	Yes	Yes	No	68.1
8.		Yes	Yes	Yes	Yes	67.3

Although, it is universally accepted that if the gap between characters is reduced more than 6 pixels, humans are not able to identify the characters from the image.

2. Rotation : We have taken the CAPTCHA schemes in which the characters are rotated clockwise/ anti-clockwise with an angle of 10, 20, 30, 45, and 60 degrees. Characters that are rotated less than 45 degrees are accurately (100.0%) recognized by the solver. However, there is decrease in AR when character is rotated with 45 degrees and 60 degrees. Generally, rotation at greater angles is discouraged due to bad user experience.

3. Distortion : Distorted characters can definitely confuse the bot as two different characters my look very much similar. And this confusion could be enhanced with the presence of occluding lines. For example, "O"(English Alphabet) and "0" (Arabic Numerical) look very similar when they are distorted. The developed solver correctly labels the distorted CAPTCHA schemes with reasonable AR of 75.2%. This shows that distortion cannot alone defeat the solver.

4. Waving : We also included the CAPTCHA schemes with various waving degrees. The solver was able to label 95.0% of the CAPTCHAs correctly. Sometimes, solver finds it difficult to detect the difference between '0' and 'O' and '1' and 'I' when waved at higher angles.

In the Table 4, we evaluate how different security features associated with CAPTCHA schemes affect the effectiveness of the solver. The literature shows that there is a decrease in AR if we include more and more security features. But proposed CAPTCHA solver maintained a high AR even when all the security features were present. CAPTCHA schemes with overlapping and rotation as security features are solved with AR of 78.5%. CAPTCHA schemes with rotation and waving as security features are solved with AR of 87.5%. Table 4 shows that when all the security features are present, the AR is 67.3%. We can conclude from the Table 4 that distortion has more negative effect on AR than other security features.

5.3. Comparison with previous work

In this section, we discuss the performance of the proposed model with the models used in literature. We compare the developed approach against 5 state-of-the-art attacks [7–11] on various CAPTCHA schemes from Microsoft, Wikipedia, Yahoo!, eBay, Weibo, etc. websites. In order to impart a fair comparison, we used the same datasets or CAPTCHA schemes used by these mentioned approaches. We tried to include CAPTCHA schemes on which these models were tested. For the same generation of different obsolete CAPTCHA schemes, we used python CAPTCHA generation tools. We produced CAPTCHA that look very similar to CAPTCHA used in state-of-the-art attack models. For each of the presented CAPTCHA schemes, we collected 1000 samples for testing. The results for comparison of the proposed approach with the approach used in [7–11] are shown in Tables 5, 6, 7, 8 and 9 respectively. Table 5 shows that the proposed approach achieved AR of 99.8% on the CAPTCHA of Reddit against 42.0% [7]. Table 6 shows that the proposed approach solved the CAPTCHA schemes of Yahoo! with AR of 55.0%. It can be seen that the AR achieved for the same CAPTCHA scheme by the technique proposed in [8] is 5.2% which increased to 55.0% using the proposed approach in this paper. The AR of the Microsoft CAPTCHA schemes in Table 6 increased from 16.2% to 60.0% which is huge. Table 7 shows that we have improved the AR from 5.3% [9] to 65.0% in case of IBM CAPTCHA schemes. In Table 8, Paypal CAPTCHA schemes which contains written text in the background frame is solved with an AR of 87.2% using the approach proposed by the authors. We have received the AR of 100.0% on the CAPTCHA scheme of Blizzard (Table 5) and eBay (Table 6). Tables 5–7 show that the proposed approach has led to a lot of improvement in the AR of Wikipedia CAPTCHA schemes as compared to the AR achieved by [7–9] respectively. In Table 9 which compares the discussed approach with latest study [11], we improved the AR of Weibo and Gregwar CAPTCHA schemes from 92.6% to 96.54% and 54.20% to 63.18% respectively. The approach discussed in this paper outperforms the approach discussed by [7–11] in the case of CAPTCHA schemes of eBay, Wikipedia, Microsoft, CNN, PayPal, etc. by a huge margin.

Table 5
Comparison of the Proposed approach with the approach discussed in [7].

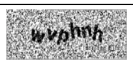
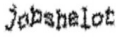
Company	Scheme	Approach in [7] AR (%)	Proposed approach AR (%)
CAPTCHA.net		73.0	98.0
Blizzard		70.0	100.0
Reddit		42.0	99.8
eBay		43.0	100.0
Wikipedia		25.0	83.0

Table 6
Comparison of the Proposed approach with the approach discussed in [8].

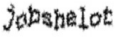
Company	Scheme	Approach in [8] AR (%)	Proposed approach AR (%)
Yahoo!		5.2	55.0
Amazon		25.8	83.0
Microsoft		16.2	60.0
eBay		58.8	100.0
Wikipedia		23.8	83.0

Table 7
Comparison of the Proposed approach with the approach discussed in [9].

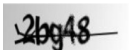
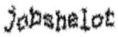
Company	Scheme	Approach in [9] AR (%)	Proposed approach AR (%)
IBM		5.3	65.0
reCAPTCHA		22.7	75.2
CNN		51.1	57.0
eBay		51.4	100.0
Wikipedia		28.3	83.0

Table 8
Comparison of the Proposed approach with the approach discussed in [10].

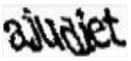
Company	Scheme	Approach in [10] AR (%)	Proposed approach AR (%)
Yahoo!		57.4	65.0
reCAPTCHA		66.6	75.2
PayPal		57.1	87.2

Table 9
Comparison of the Proposed approach with the approach discussed in [11].

Company	Scheme	Approach in [11] AR (%)	Proposed approach AR (%)
Weibo		92.68	96.54
Gregwar		54.20	63.18

6. Conclusions and future works

This paper has delivered very simple and generic algorithm for solving or breaking CAPTCHA schemes. The proposed approach uses Depth First search for extracting the characters from the CAPTCHA schemes and Convolutional Neural Network for recognizing these extracted characters. The proposed approach requires very little or no human interference when targeting a new CAPTCHA set. This approach is very generic and can be easily modified to adjust with constantly evolving CAPTCHAs. This approach is validated with 30000+ text CAPTCHA (simple, medium and complex) schemes collected from 30+ websites used by popular Indian government and MNCs which includes security features like overlapping, rotation, waving and distortion. The results show that these advanced security features only introduce difficulties for a legitimate user but fail to stop cyber attack done by advanced machine learning algorithms such as the approach discussed in this paper. Our proposed approach raises questions for the popular websites (IRCTC, UIDAI, SBI Online, AliBaba) which are employing the use of CAPTCHAs on how to redesign the existing CAPTCHAs for ensuring security in future. The performance of the proposed approach can be summarized as below:

1. We have achieved an average accuracy rate of more than 92.0% in detecting CAPTCHAs.
2. We have achieved an average accuracy rate of 98.0%, 90.6%, and 71.9% for easy, medium, and complex CAPTCHAs.
3. On an average, model solved simple CAPTCHAs like IRCTC within 30.20 ms with average accuracy rate of 98.0%. The medium complexity CAPTCHAs like UIDAI, Digital India, Delhi Police etc. are solved between 40.00 to 45.00 ms with average accuracy rate of 90.6%. The difficult ones like AliBaba, SBI Online, CAT 2019, India.gov.in etc. are solved within 45.00 to 50.00 ms with average accuracy rate of 71.9%. There is always room for further improvement and possible advancement. This model can be extended to recognize the letters of different languages by collecting the dataset and training the CNN model on it. Treating all the words in a single sequence or first separating the words and then recognizing the individual words can solve multi-word CAPTCHA schemes. We can pre-calculate characters present to deal with variable number of characters containing the CAPTCHA schemes. It can also be extended to video and audio-based CAPTCHA schemes by segmenting the video frames into images and then recognize the characters from individual images. Hence, the insights provided in this paper can help security experts to revisit the viability of the text CAPTCHAs.

CRediT authorship contribution statement

Aditya Atri: Writing – original draft, Writing – review & editing, Conceptualization, Data curation, Validation. **Ankita Bansal:** Writing – original draft, Writing – review & editing, Conceptualization, Data curation, Validation. **Manju Khari:** Writing – review & editing, Conceptualization, Formal analysis, Supervision. **S. Vimal:** Writing – review & editing, Conceptualization, Formal analysis, Data curation, 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.

References

[1] Zhao Huawei, Bai Peidong, Peng Yun, Xu Ruzhi. Efficient key management scheme for health blockchain. 3, CAAI Transactions on Intelligence Technology; 2018, p. 114–8.
[2] Kumari Suchi, Namasudra Suyel. System reliability evaluation using budget constrained real d-MC search. Comput Commun 2021;171.
[3] Alguliyev Rasim M, Sukhostat yudmila V. Efficient algorithm for big data clustering on single machine. 5, CAAI Transactions on Intelligence Technology; 2020, p. 9–14.
[4] Jiang Jingyan, Hu Liang. Decentralised federated learning with adaptive partial gradient aggregation. 5, CAAI Transactions on Intelligence Technology; 2020, p. 230–6.
[5] Coates Adam, Carpenter Blake, Case Carl, Satheesh Sanjeev, Suresh Bipin, Wang Tao, et al. Text detection and character recognition in scene images with unsupervised feature learning. In: Proceedings of the 2011 international conference on document analysis and recognition. ICDAR '11, USA: IEEE Computer Society; 2011, p. 440–5.
[6] Chow Richard, Golle Philippe, Jakobsson Markus, Wang Lusha, Wang XiaoFeng. Making CAPTCHAs clickable. In: Proceedings of the 9th Workshop on Mobile Computing Systems and Applications. HotMobile '08, New York, NY, USA: Association for Computing Machinery; 2008, p. 91–4.
[7] Bursztein Elie, Aigrain Jonathan, Moscicki Angelika, Mitchell John C. The end is nigh: Generic solving of text-based CAPTCHAs. In: Proceedings of the 8th USENIX conference on offensive technologies. WOOT'14, USA: USENIX Association; 2014, p. 3.
[8] Gao Haichang, Yan Jeff, Cao Fang, Zhang Zhengya, Lei Lei, Tang Mengyun, et al. A simple generic attack on text captchas. In: Network and distributed system security symposium. 2016.
[9] George Dileep, Lehrach Wolfgang, Kansky Ken, Lázaro-Gredilla Miguel, Laan Christopher, Marthi Bhaskara, et al. A generative vision model that trains with high data efficiency and breaks text-based CAPTCHAs. Science 2017;358:eaag2612.
[10] Goodfellow Ian J, Pouget-Abadie Jean, Mirza Mehdi, Xu Bing, Warde-Farley David, Ozair Sherjil, et al. Generative adversarial nets. In: Proceedings of the 27th international conference on neural information processing systems - volume 2. NIPS'14, Cambridge, MA, USA: MIT Press; 2014, p. 2672–80.
[11] Thobhani Alaa, Gao Mingsheng, Hawbani Ammar, Ali Safwan Taher Mohammed, Abdussalam Amr. CAPTCHA Recognition using deep learning with attached binary images. Electronics 2020;9:1522.
[12] Krol Kat, Parkin Simon, Sasse Angela. Better the devil you know: A user study of two CAPTCHAs and a possible replacement technology. In: Workshop on useable security; 2016.
[13] Mori Greg, Malik Jitendra. Recognizing objects in adversarial clutter: breaking a visual CAPTCHA. In: Proceedings of the 2003 IEEE computer society conference on computer vision and pattern recognitio; 2003. p. 134–41.

[14] Yan Jeff, El Ahmad Ahmad Salah. A low-cost attack on a microsoft captcha. In: Proceedings of the 15th ACM Conference on Computer and Communications Security. CCS '08, New York, NY, USA: Association for Computing Machinery; 2008, p. 543–54.
[15] Yan Jeff, Ahmad Ahmad Salah El. Breaking visual CAPTCHAs with naive pattern recognition algorithms. In: Twenty-third annual computer security applications conference (ACSAC 2007); 2007. p. 279–91.
[16] Gao Haichang, Tang Mengyun, Liu Yi, Zhang Ping, Liu Xiyang. Research on the security of microsoft's two-layer captcha. IEEE Trans Inf Forensics Secur 2017;12:1671–85.
[17] Ye Guixin, Tang Zhanyong, Fang Dingyi, Zhu Zhanxing, Feng Yansong, Xu Pengfei, et al. Yet another text captcha solver:A generative adversarial network based approach. In: Proceedings of the 2018 ACM SIGSAC conference on computer and communications security; 2018.
[18] Sano Shotaro, Otsuka Takuma, Ityyama Katsutoshi, Okuno Hiroshi G. HMM-based attacks on google's ReCAPTCHA with continuous visual and audio symbols. J Inform Process 2015;23:814–26.
[19] Sivakorn Suphannee, Polakis Jason, Keromytis Angelos D. I ' m not a human : Breaking the Google reCAPTCHA. 2016.
[20] Namasudra Suyel. Fast and secure data accessing by using DNA computing for the cloud environment. IEEE Trans Serv Comput 2020.
[21] Xu Weilin, Qi Yanjun, Evans David. Automatically evading classifiers: A case study on PDF malware classifiers. In: Network and distributed system security symposium; 2016.
[22] Goodfellow Ian, Shlens Jonathon, Szegedy Christian. Explaining and harnessing adversarial examples. 2014, ArXiv:1412.6572.
[23] Zhang Yang, Gao Haichang, Pei Ge, Luo Sainan, Chang Guoqin, Cheng Nuo. A survey of research on CAPTCHA designing and breaking techniques. 2019, p. 75–84.
[24] Namasudra S, Dhamodharavadhani S, Rathipriya R. Nonlinear neural network based forecasting model for predicting COVID-19 cases. Neural Process Lett 2021.
[25] Krizhevsky Alex, Sutskever Ilya, Hinton Geoffrey. ImageNet Classification with deep convolutional neural networks. Neural Inf Process Syst 2012;25.

Aditya Atri has pursued Bachelor of Engineering in Information Technology (2016–2020) from Netaji Subhas Institute of Technology (NSIT), Delhi. His research interests are big data, machine learning and distributed computing.

Ankita Bansal is an Assistant Professor at the Department of Information Technology, Netaji Subhas University of Technology, India. Prior to that, she worked as research scholar at Delhi Technological University, India. She received her master's and doctoral degree in Computer Science from Delhi College of Engineering. Her research interests are software quality, soft computing, database management, machine learning and meta heuristic models.

Manju Khari is an Associate Professor in School of Computer and System Sciences, Jawaharlal Nehru University, India. Earlier she served as Assistant Professor in Ambedkar Institute of Advanced Communication Technology and Research. She holds a Ph.D. in Computer Science Engineering from National Institute of Technology Patna. Her research interests are software testing, information security, optimization, image processing and machine learning.

S. Vimal is working in Department of Artificial Intelligence and Data Science, Ramco Institute of Technology, Tamilnadu. Earlier he served in Department of Information Technology, National Engineering College, Kovilpatti, Tamilnadu. He holds a Ph.D. in Information and Communication Engineering from Anna University Chennai. His areas of interest include Game Modeling, Artificial Intelligence, Cognitive radio networks, network security, machine learning and big data analytics.

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)
Pune, India. Oct 29-30, 2021

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON) | 978-1-6654-2503-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/SMARTGENCON51891.2021.9645786

Devanagari Handwritten Character Recognition using CNN as Feature Extractor

Shilpa Kaur Manocha
Dept. of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
shilpamanocha03@gmail.com

Piyush Tewari
Dept. of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
piyushtewari@dtu.ac.in

***Abstract*— This paper proposes the use of Convolutional Neural Networks as feature extractor for extraction of features from handwritten Devanagari characters. For classification, classifiers employed are SVM (Linear, Polynomial and RBF), KNN, RF, DT, MLP and XGB. Use of CNN model for feature extraction eliminates the need of handcrafted features by traditional pattern recognition methods. XGB Classifier, which is a very recent technique, has also been explored (which has not been done previously to the best of our knowledge). The experiments with these eight techniques have been done on the DHCD dataset proposed in year 2015. Use of CNN proved to be very effective for Devanagari characters recognition as all the models achieved recognition accuracy of over 92% and total training time including feature extraction and classification did not exceed a total of 12.16 minutes.**

Keywords*— *CNN, SVM, KNN, DT, RF, OCR, HCR, XGB, MLP

I. INTRODUCTION

Devanagari handwriting recognition has been gaining attention by researchers in recent years. The script serves as a base for many popular scripts like Hindi, Marathi, Sanskrit etc. Unfortunately, there are no robust commercial systems available for recognition of Devanagari handwritten characters. Owing to difficulties like the cursive nature of script, many characters differing only in minor structural details, different handwriting styles and non-availability of standard labelled datasets, the recognition task becomes tedious. Development of robust OCR[1] systems for recognition of handwritten characters of Devanagari script[2] will enable us to preserve old manuscripts by converting the physical files to digital formats.

For conversion of handwritten characters from physical documents and images to text, the first step is image pre-processing [3] that includes:- binarization, skew/slant/rotation removal, image thresholding and noise removal. The next step is feature extraction i.e. instead of feeding all raw pixels as input to the classifier for classification, several methods are employed to detect important features from the images. For textual data, relevant features are extracted using horizontal and vertical projections, crossings, distances, shape number, Histogram of gradient etc. Selection of features that can best describe the dataset and their extraction is a crucial step and affects the final recognition accuracy. Finally in the last step, a classifier takes the combination of these extracted features as inputs and classifies the characters into appropriate labels.

In this paper, a novel method for of the recognition of handwritten Devanagari characters has been presented. Feature extraction using Convolutional Neural Network model has been utilised that eliminates the need of explicit handcrafted feature selection and extraction. Manual selection of best features and extraction using various methods is a tedious task and doesn't always result in best accuracy. Main contributions of this paper include: -

1) Hybrid recognition method for automatic feature selection and classification of handwritten Devanagari characters. CNN model is used as feature extractor instead of traditional pattern recognition techniques and use of SVM, KNN, DT, RF, MLP, XGB as classifiers.
2) Characters classification with the use of XGB classifier: both with the input as raw pixels and with CNN as feature extractor.

The paper is arranged in following sections: - Section 1 gives Introduction, Section 2 is dedicated to literature review, Section 3 gives brief explanation of classification techniques, Section 4 describes about the proposed system architecture, Section 5 is dedicated to Results and Discussion and Section 6 provides conclusion of the experiments performed.

II. LITERATURE REVIEW

In [4], the authors made use of discrete cosine transform for feature extraction and Neural network based classification for recognition of Devanagari characters. The highest recognition accuracy obtained was 94.89%. In [5], the authors classified 1000 Hindi characters by calculating the histogram of projections and then classifying the characters with artificial neural network. The classification accuracy obtained was 90%. In [6], the authors have recognised Devanagari script characters by finding the gradient and curvature of characters and then classifying them using two classifiers namely MQDF and SVM. The highest recognition accuracy obtained was 95.13%. In [7], the authors employed GLAC feature extraction method and SVM classifier for recognition of Devanagari handwritten characters. Accuracies reported on the datasets: V2DMDCHAR and ISIDCHAR are 95.21% and 93.21% respectively. In [8], the authors classified Devanagari characters by combining wavelet based feature extraction and BPNN classifier. The highest accuracy achieved was 70%. In [9], the author proposed a deep convolutional NN for recognition of Devanagari handwritten alphabets and digits. The model obtained an accuracy of 98.47%. In [10], the authors developed a recognition system for Devanagari script by evaluating various structural and geometrical features and performing classification using MLP classifier.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:59:04 UTC from IEEE Xplore. Restrictions apply.

Maximum classification accuracy achieved was 82.7%. In [11], the authors developed a recognition system for Devanagari script by evaluating various structural features and performing classification using FFBPN, CFBPN and EBPN classifiers. Accuracy rates obtained were 97.20%, 97.46% and 98.10% respectively. In [12], the authors use of AlexNet model that has been pertained on the ImageNet dataset, for building a recognition system for Devanagari script characters. The results demonstrated that in 3 epochs only, the accuracy rate went above 90%. The highest validation accuracy and test accuracy attained were 94.49 % and 95.46% respectively. In [13], the authors have proposed a deep CNN model for classification of Devanagari handwritten characters. The model was able to identify the characters with 93.73% accuracy on a small database of size 5484.

III. Related Theory

A. Support Vector Machine(SVM)

Support Vector Machine [14] is a supervised machine learning technique used for classification and regression tasks. SVM is a popular and widely used classifier for image classification tasks. It distinguishes the sample points by identifying a hyperplane that separates the classes. The sample points that are nearest to the hyperplane are defined as "support vectors". The classifier aims to maximise the margin (difference between support vectors) on each side of hyperplane. The classifier takes 2 inputs: one array of shape (number_of_samples, number_of_features) and the other array containing class labels, either as numerical or string format, of shape (number_of_samples). A kernel function must be specified which can be linear, polynomial, radial basis function or any other custom function.

1) Linear Kernel: - {x, x'}
2) Polynomial (Poly): - {γ (x,x') + r) ^ d , where γ- , d is degree, r- coefficient.
3) Radial Basis Function (RBF): $(-\gamma|x-x'|2)$, where γ is gamma (> 0).

The classifier has following advantages :- powerful in high dimensional spaces, viable in cases where number of dimensions are more than the number of samples, is memory effective as it only focuses on a subset of training points i.e. the support vectors and is versatile i.e. different kernel functions can be used.

The disadvantages of the classifier are : overfitting may occur in cases where number of feature elements are larger than the number of sample points, selection of kernels is difficult and calculation of scores & probabilities is done using expensive k-fold cross-validation.

B. K-Nearest Neighbours (KNN)

KNN [15] is a supervised learning method. The idea behind this algorithm is that similar things occur in close proximity. Closeness between points is measured using distance between the points .e.g. by using Euclidean distance. It finds distance between the query and all data points , and selects a specified number of samples (k) closest data points from query. The query is classified to the class label that is the most repeating in the chosen k points group. The advantages of using KNN are:- easy and simple to implement, no need of tuning parameters, can be employed for classification as well as regression tasks. The disadvantage of algorithm lies in the slow classification time for large number of samples/classes. KNN algorithm is a popular choice in recommender systems, as they work on principle of close data points.

C. Decision Tree(DT)

Decision Tree [16] is a powerful classification and prediction tool that learns using an inductive approach. It has a tree structure similar to that of a flowchart. In a decision tree, there are several nodes ,that present a decision condition, starting from parent node at the top to the terminal node as the final class label. At each node, the data is classified based on a test on the attribute and each branch represents the result of the test/decision. A DT is built by recursively splitting the data samples into subsets based on the attribute tests at each node. The splitting is stopped when we have the final predicted classes at the final nodes. An optimal DT is one that is able to represent all data points with least number of node levels. Pruning methods are used to drop nodes that are not critical for classification ,hence reducing size of DT. Use of DT has following advantages:- they can handle both categorical and numerical values, no prior domain knowledge requirement, can be visualised , results can be interpreted and analysed to understand critical decision rules. DT have following limitations:- may lead to overfitting on data by creating complex tree, problem of variance i.e. change in tree with even slight changes in data points, no guarantee of optimal tree by use of greedy algorithms and it results in a biased tree if some classes in the dataset dominate.

D. Random Forest (RF)

Random Forest Classifier [17] is a supervised learning algorithm that combines decision tree classifier with the bagging method. A forest is a set of decision trees, that collectively vote for the correct label. The model builds multiple trees by randomly a subset of features for each tree. Thus multiple trees of reduced depth are built and overfitting is reduced. The results from subtrees are averaged to reach to final class label. The classifier also makes it possible to measure the importance of features on a scale of 0-1. Important features are those having higher importance value. The classifier doesn't require much hyper-parameter tuning, is simple and easy to build, reduces chances of overfitting and mostly results in good accuracy for both classification as well as regression tasks. The training of model doesn't need much time but prediction takes some time. In case of large number of trees, model gives higher accuracy but the computations become slow.

E. Multi Layer Perceptron (MLP)

Multi-layer perceptron classifier [18] is a deep, artificial neural network (ANN) .It has 3 types of layers- 1 input layer, multiple hidden layers & one output layer. The input layers are connected with output layers as directed graph. The classifier trains by method of back-propagation. It can be used as a classifier or a regressor. The model minimises cross entropy loss function for classification. For classification into multiple classes, 'SoftMax' function is used for adding non-linearity. The advantages of using MLP are:- it can understand non-linear models and can learn models in real time. The weaknesses include: requirement of fine tuning of hyper parameters like number of hidden units, number of layers and iterations, difference in validation accuracy with different random weight initialisations due to non-convex

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:59:04 UTC from IEEE Xplore. Restrictions apply.

loss function of hidden layers. The model is also sensitive to feature scaling.

F. Extreme Boost Gradient (XGB)

XGB [19] Algorithm can be used for both classification and regression tasks. Boosting is an iterative approach to building ensemble of trees. In XGB classifier, gradient boosting is done by building trees iteratively. The trees try to reduce/correct residual errors from previous trees. In other ensemble methods, the models end up learning same mistakes. The building of large number of sequential trees creates a complicated model which may lead to overfilling. But the model provides parameter tuning methods to prevent it.

G. Convolutional Neural Network (CNN)

Convolutional NNs [20] are popularly used for image processing & computer vision tasks. They consist of combination of convolution and pooling layers followed by fully connected layers. The convolutional layers consist of filters and are convolved with input images to extract relevant features. The pooling layer reduces the dimensions of feature maps by performing average or maximum pooling over each area in the image. The fully connected layers activate the neurons and classifies the input features into appropriate classes.

IV. Proposed System Architecture

A. Framework

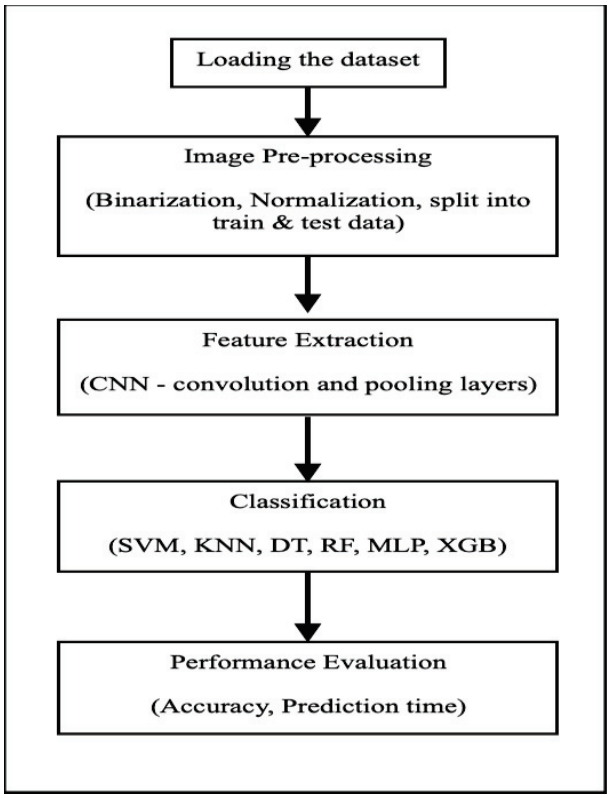


Fig. 1. Flowchart of proposed recognition method

Step1) Loading the dataset: The dataset consists of isolated characters of each character class in separate folders. The raw pixel information is available in a csv file. Each character is of size 32*32 i.e., having 1024-pixel values. The character labels have also been mentioned for all characters. Input features (1024) are read into one array and character labels in another array.

Step2) Image Pre-processing: The pixels contain information of the gray scaled images. The images are converted to binary. Normalisation is performed for all features by dividing them by 255(highest pixel value in black and white image). Thus, we get features normalised to a scale of 1. The features and character label arrays are split into training and testing set.

Step 3) Feature Extraction using CNN: CNN automatically extracts relevant features from the data fed. Convolution layers are used to create feature maps from training samples and pooling layers are used for reducing the size of feature maps generated. Instead, or connected the last fully connected dense layer for classification, the features are obtained in an array and fed to different classifiers as input. The model is trained for 3 epochs to learn the features from the training samples. Figure 2 shows the CNN model architecture used for the purpose of feature extraction.

Model: "sequential_1"

Layer (type)	Output Shape	Param #
conv2d_4 (Conv2D)	(None, 30, 30, 32)	320
conv2d_5 (Conv2D)	(None, 28, 28, 64)	18496
max_pooling2d_2 (MaxPooling2	(None, 14, 14, 64)	0
conv2d_6 (Conv2D)	(None, 12, 12, 64)	36928
conv2d_7 (Conv2D)	(None, 10, 10, 64)	36928
max_pooling2d_3 (MaxPooling2	(None, 5, 5, 64)	0
dropout_1 (Dropout)	(None, 5, 5, 64)	0
flatten_3 (Flatten)	(None, 1600)	0
dense_3 (Dense)	(None, 128)	204928
dense_4 (Dense)	(None, 64)	8256
dense_5 (Dense)	(None, 46)	2990

Fig. 2. CNN Model Architecture used for feature extraction

Step 4) Classification: Features extracted from last layer of CNN model are fed to various classifiers and test samples are applied for prediction. Classifiers are fed with two arrays, X - input features obtained from last/Dense layer of CNN and y-class labels for training and then tested on test samples. Total number of character classes are 46. Default parameter values have been used for each classifier.

Step 5) Performance Evaluation: For performance analysis , classification accuracy and prediction time are noted for all classifier models.

V. Results And Discussion

A. Dataset Used

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:59:04 UTC from IEEE Xplore. Restrictions apply.

For Experimental Analysis, we have made use of the DHCD [21] dataset, proposed by Acharya et. al, in 2015. The dataset includes 46 handwritten Devanagari characters - 36 alphabets and 10 numerals. Each character label has 2k image samples. i.e., a total of 92k samples. The dataset is divided into two parts: training set -to train and extract the features from CNN and testing set - fed to classifiers for testing the performance of classifier. We have considered 3 cases for dataset splitting. Case 1) 70:30, Case 2) 75:25, Case 3) 80:20.

Fig. 3. Handwritten characters from the DHCD dataset

B. Evaluation Parameters

1) Accuracy: It is the ratio of correct predictions made by the classifier to the total number of test samples fed. High accuracy percentage is the desired outcome.

2) Prediction Time: The time required by the classifier to segregate the test samples into different classes, after they have been fed with input features to be classified. Lowest prediction time for correctly identifying the classes is the desired outcome.

TABLE 3. Recognition Results of different classifiers on DHCD dataset.

Model / System	**Case 1: Dataset Split - 70:30**		**Case 2: Dataset Split - 75:25**		**Case 3: Dataset Split - 80:20**	
	Accuracy(in %)	**Prediction Time (in seconds)**	**Accuracy(in %)**	**Prediction Time (in seconds)**	**Accuracy(in %)**	**Prediction Time (in seconds)**
CNN + KNN	98.73	21.35	98.91	30.84	98.99	25.60
CNN + SVM Linear	98.41	**10.78**	**98.40**	**10.80**	98.87	12.03
CNN + SVM Poly	98.36	20.82	98.63	20.79	98.74	22.20
CNN + SVM RBF	**98.88**	23.78	99.01	23.21	99.13	26.60
CNN + MLP	98.55	22.84	98.69	17.73	98.86	30.03
CNN + DT	93.54	**2.95**	**93.85**	**2.79**	94.41	**3.43**
CNN + RF	98.57	20.21	98.84	20.20	98.7	23.88
CNN + XGB	98.15	169.92	98.37	160.59	98.42	198.51
Feature Extraction time taken by CNN (in seconds)	497.36		467.68		532.48	

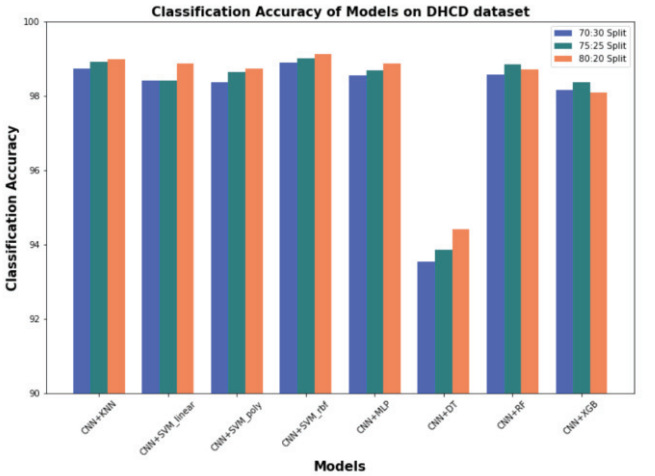


Fig. 4. Comparison of Classification accuracy of models on the DHCD [21] Dataset

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:59:04 UTC from IEEE Xplore. Restrictions apply.

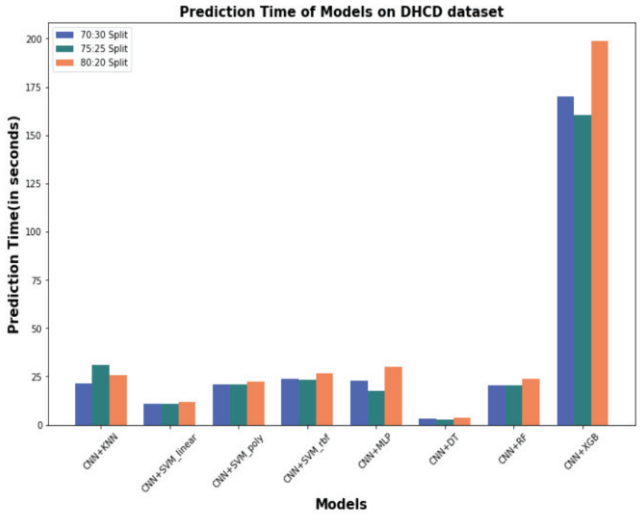


Fig. 5. Comparison of prediction time required by various models on the DHCD [21] Dataset

VI. CONCLUSION

It has been observed that CNN is able to identify relevant features from images and thus results in computational cost saving as compared to traditional pattern recognition techniques for feature extraction. It also saves the hassle and struggle of analysing and identifying best features for the image dataset and then applying various techniques or combination of feature extraction techniques to get appropriate feature maps. Use of CNN as feature extractor reduced the training and prediction time significantly for all classifiers(in the range of few seconds), along with achieving accuracy rates as high as 99% (CNN + SVM_RBF). The Devanagari characters have also been classified using XGB classifier which gives approx. 90% accuracy when trained with raw pixel values which is unusually remarkable. Use of CNN as feature extraction and XGB as classifier further improved the accuracy to 98% with average prediction time being only 176.34 seconds.

REFERENCES

[1] Mithe, R., Indalkar, S., & Divekar, N., "Optical character recognition", International journal of recent technology and engineering (IJRTE), 2(1), 72-75,2013. OCR

[2] Bright, W.,. "The devanagari script.", The world's writing systems, pp.384-390,1996. Script

[3] Kumar, G., Bhatia, P.K. and Banger, I. "Analytical review of preprocessing techniques for offline handwritten character recognition.", International Journal of Advances in Engineering Sciences, 3(3), pp.14-22, 2013.

[4] Kubatur, S., Sid-Ahmed, M., & Ahmadi, M. (2012). A neural network approach to online devanagari handwritten character recognition. In IEEE International Conference on High Performance Computing and Simulation (pp. 209–214). Madrid: IEEE Press.

[5] Yadav,D.,Sanchez-Cuadrado,S.,&Morato,J.(2013).Optical character recognition for Hindi language using a neural-network approach. Journal of Information Processing Systems (JIPS), 9(1), 117–140.

[6] Pal,U.,Chanda,S.,Wakabayashi,T.,&Kimura,F.(2008).Accuracy Improvement of Devnagari character recognition combining SVM and MQDF. In 11th International Conference on Frontiers in Handwriting Recognition (pp. 367–372). Canada: L. Lam Publications.

[7] Jangid,M.,&Srivastava,S.(2014).Gradient localauto-correlation for handwritten Devanagari character recognition. In IEEE International Conference on High Performance Computing and Applications (pp. 1–5). Bhubaneswar: IEEE Press.

[8] Dixit,A.,Navghane,A.,&Dandawate,Y.(2014).Handwritten devanagari character recognition using wavelet based feature extraction and classification scheme. In: Annual IEEE India Conference (pp. 1–4). Pune: IEEE Press.

[9] Acharya,S.,Pant,A.K.,&Gyawali,P.K.(2015).Deep learing based large scale handwritten devanagari character recognition. 9th International Conference on Software, Knowledge, Information Management and Applications (pp. 1–6). Nepal: IEEE Press.

[10] Dongre,V.J.,&Mankar,V.H(2015).Devanagari offline handwritten numeral and character recognition using multiple features and neural network classifier. In 2nd International Conference on Computing for Sustainable Global Development (pp. 425–431). New Delhi: IEEE Press.

[11] Shelke, S., & Apte, S. (2016). Performance optimization and comparative analysis of neural networks for handwritten Devanagari character recognition. In IEEE International Conference on Signal and Information Processing (pp. 1–5). Vishnupuri: IEEE Press. 106 S. Singh and N. K. Garg

[12] P. K. Sonawane and S. Shelke, "Handwritten Devanagari Character Classification using Deep Learning.," 2018 International Conference on Information , Communication, Engineering and Technology (ICICET), pp. 1-4, 2018.

[13] Narang, S.R., Kumar, M. & Jindal, M.K." DeepNetDevanagari: a deep learning model for Devanagari ancient character recognition." Multimed Tools Appl 80, 20671–20686, 2021.

[14] Ramya, S., & Shama, K. Comparison of SVM Kernel effect on online handwriting recognition: A case study with Kannada script. In Data engineering and intelligent computing (pp. 75-82), 2018. Springer, Singapore.

[15] Zhang, M. L., & Zhou, Z. H. ML-KNN: A lazy learning approach to multi-label learning. Pattern recognition, 40(7), 2038-2048, 2007.

[16] Song, Y. Y., & Ying, L. U. Decision tree methods: applications for classification and prediction. Shanghai archives of psychiatry, 27(2), 130, 2015.

[17] Bernard, S., Adam, S., & Heutte, L. Using random forests for handwritten digit recognition. In Ninth international conference on document analysis and recognition (ICDAR 2007) (Vol. 2, pp. 1043-1047), 2007. IEEE.

[18] Satange, D. D., Ajmire, D. P., & Khandwani, F. I. Offline Handwritten Gujrati Numeral Recognition Using MLP Classifier. International Journal of Novel Research and Development, 3(8), 2018.

[19] Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., & Cho, H. Xgboost: extreme gradient boosting. R package version 0.4-2, 1(4), 1-4, 2015.

[20] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," International Conference on Engineering and Technology (ICET), pp. 1-6, 2017.

[21] S. Acharya, A. K. Pant and P. K. Gyawali, "Deep learning based large scale handwritten Devanagari character recognition," 2015 9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA), pp. 1-6,2015.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:59:04 UTC from IEEE Xplore. Restrictions apply.

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)
Pune, India. Oct 29-30, 2021

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON) | 978-1-6654-2503-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/SMARTGENCON51891.2021.9645839

EASE Em: Hybrid Recommendation System for Item Cold Start Problem

S. Aditya
Computer Science and Engineering
Delhi Technological University, Delhi
s.aditya.me@gmail.com

Mihir Rajora
Computer Science and Engineering
Delhi Technological University, Delhi
mihee20@gmail.com

Indu Singh
Assistant Professor
Computer Science and Engineering
Delhi Technological University, Delhi
indusingh@dtu.ac.in

***Abstract*—Recommender Systems play a very important role in online services nowadays, with the huge repository of content being generated every day. While these systems are used on large scales, they still face a lot of problems. The item cold start is one of the most famous problems in recommender systems. While systems can develop an initial image of new user profiles through various methods, new items face a problem surfacing into user recommendations. This is because they have low to almost zero interactions, and most recommender systems use methods like collaborative filtering to share similar items to similar user profiles. This is an important problem to be dealt with. Otherwise, recommendation feeds get saturated with already popular content, which harms new creators on platforms that might be creating innovative content. We propose a hybrid recommender system: EASE'Em, that seeks to solve this problem by generating cold start items to users based on the item similarities and user profiles. Our model first uses an AutoEncoder recommendation system to fill a sparse user interaction matrix to develop a better understanding of user profiles. We then project item attributes and user profiles into a common low dimensional space and then use matrix factorization techniques to generate top-N recommendations for each user present in the set. We then compare our model generations with standard item-based recommender models on a subset of the Million Song Dataset and compare their results.**

Keywords: AutoEncoders, Closed Form Optimization, Hybrid Models, Item Cold Start Problem, Neighbourhood-Approach, Recommender System

I. Introduction

With increasing popularity in the usage of deep learning techniques for problems ranging from recommendation to classification, almost all multimedia recommender systems have been aimed towards methods such as collaborative filtering[1].

In this paper, we try to implement a similar approach for the generating multimedia recommendations for the users, but without any hidden layers, i.e., we define a linear model with an input and an output layer. We further try to optimise the learning functions by deriving closed-form solutions for the minimum value of the loss functions by making use of Lagrangians as defined in [2]. The advantage of deriving a closed-form solution for the training of the model is that once the dataset is loaded into memory, the running time for training the model is several orders less than that for training a model based on collaborative filtering and neural networks.

Further, since many recommendation problems revolved around making predictions based on past user interaction history, it proved to be detrimental to songs that were new to the system and had no previous user interactions, since no model that relied solely on user interactions could recommend that item. Further, in case of items where although a user interaction exists, but the disparity among the items receiving a high number of interactions with the items having low interactions, only the *popular songs* get recommended. This has been referred to in [3] as Popularity Bias and the overall problem involving such items in the system is referred to as the *Item Cold Start Problem*.

We try to solve the *Item Cold Start Problem* by making use of the metadata information associated with the multimedia. This implies that given the feature description of items and the user interaction of their user interaction, we aim to generate a vector of users associated with each item that are probable to consume that media. By doing that, we aim to solve the *Item Cold Start Problem*.

Hence, we propose a hybrid multimedia recommender system that efficiently generates recommendations for the user based on past interactions of all users in the system as well as suggests certain items that are *new* to the system, based on the item features.

Our paper can be outlined as follows:

1) We define a model for recommending items based on past user-item interactions using a linear model without any hidden layers, and then solving the optimisation problem using a closed-form solution
2) We define a second model for recommending songs with no past user interactions, and overcoming the *Item Cold Start Problem*.
3) We try to combine the above-mentioned models into a single recommendation system that generates a playlist for the user which suits their taste
4) We conduct experimental analysis of our model as well as a comparative study with **6** state-of-the-art recommender models.

II. Related Work

Even in today's time, there exist many problems in recommender systems today. In Schedl et al[4], the main problems in recommender systems lay within the field of Cold Start problems. While user cold start problems are dealt with in a

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:18:07 UTC from IEEE Xplore. Restrictions apply.

variety of ways, not much innovation has been done around the issue of item cold start area. The most successful results show hybridized approaches to using content-based recommender systems alongside collaborative filtering. Another challenge is the evaluation of recommender systems. Since recommendations don't have a direct ground truth to evaluate with, numerical evaluation of recommendation systems remains somewhat subjective to the definition of the evaluation parameters.

Previously, people have tried hybrid approaches for combining user interaction with item features to get more understandable recommendations. Domingues et al[5] designed a real-time recommender system that specialised on the very Long Tail of music content (content that was interacted with less, or pseudo-cold start items). They do so by evaluating on both item similarity models and user interaction models.

Other approaches have been used to develop more context on user profiles based upon their interaction with different types of media. Chiliguano and Fazekas[6] CNNs to project audio segment into an n-dimensional vector space, and use the EDA approach for developing user profiles. They use both these inputs to develop a hybrid recommender system for recommendations based on the content-based recommendation system.

Similar work has been done previously on integrating user-profiles and item feature attributes. Ning and Karypis[7] in their paper titled SLIM, solved an l1-norm and l2-norm regularization problem to generate top N recommendations through their model. They enforced non-negativity and zero-diagonal conditions on their weight matrices to ensure better output generations. While their work showed competitive accuracy in relation to other linear recommendation models, their SLIM model, like most machine learning models, was computationally expensive to train.

III. EASE'EM Recommender System

We propose a recommender system called **EASE'Em** which is a hybrid recommendation generator that, apart from recommending songs that already exist in the system, also simultaneously recommends new songs as defined in the *Item Cold Start Problem*. We train our model on the Subset of the Million Song Dataset[8] which is an open sourced collection of audio features and metadata for a million contemporary music items.

A. Model Architecture

The suggested model is designed to work in such a way that given the information of past user interactions with each song and the description and metadata associated with each song, it outputs a playlist of variable lengths to each user.

The generation of playlist is divided into two parts on the basis of how the recommendations are being generated

- **AutoEncoder Recommendation Constructor or ARC** based on the User Item Interactions, and
- **Common Dimension Projector or CDP** based primarily on the Item Features and Metadata

For **ARC**, we have the User Interaction Data in the form of a sparse matrix $X \in \mathbb{R}^{|U|\times|I|}$, where $|U|$ is the magnitude of the set of users, and $|I|$ is the magnitude of the set of all items (songs) as shown in Figure 1.

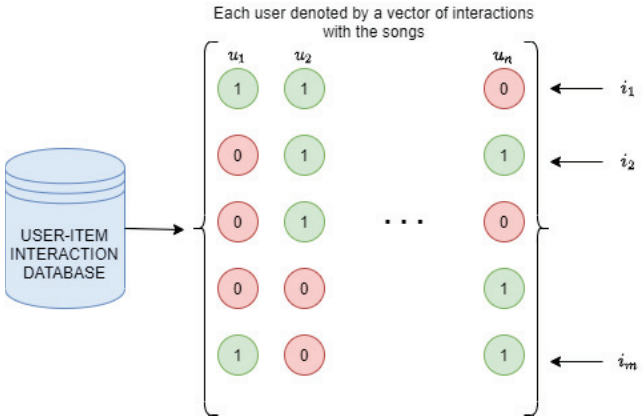


Fig. 1. User Item Interactions Data

The elements of X are defined as given in the following function,

$$X_{ij} = \begin{cases} 0 & \text{user j has not interacted with item i} \\ 1 & \text{user j has interacted with item i} \end{cases} \quad (1)$$

The weight matrix that linearly calculates the output matrix given the input matrix is given by $B \in \mathbb{R}^{|I|\times|I|}$

Similar to other neighbourhood based approaches as explored in the previous sections, we set the self-similarity of an item in the input layer with itself in the output layer to 0. This is done so that the input cannot force the model to generalise when reproduced as an output. To enforce the same the diagonal elements of B are set to 0, i.e., $\text{diag}(B) = 0$.

Thus for ARC, the score S associated with each item $i \in I$ for a given user $u \in U$ is given by the dot product of the input matrix X and the weight matrix B, represented as follows:

$$S = X \cdot B \quad (2)$$

Equation 2 can alternatively be expressed as -

$$S_{u,j} = X_{u,\cdot} \cdot B_{\cdot,j} \quad (3)$$

where $X_{u,\cdot}$ refers to the row of user u and $B_{\cdot,j}$ refers to the column of item j

For the **CDP**, we have been given a new item and its corresponding features. For each such new item we have to return a vector space of users who are likely to consume that media.

Formally, the model and its workings can be defined as follows: Given a input vector space $X_s \in \mathbb{R}^{n\times m}$ where n corresponds to magnitude of the set of training items and m refers to the number of features for each item that the model is trained on. We have one additional input matrix $X_u \in \mathbb{R}^{n\times u}$ where u indicates the number of users. Further, each element of the matrix X_u, given by $X_u(i,j)$ indicates whether user j has shown an interest towards item i.

The aim of the model is, during test time, to predict for each new item q with feature set $q_s \in \mathbb{R}^{l\times m}$ a vector of scores

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:18:07 UTC from IEEE Xplore. Restrictions apply.

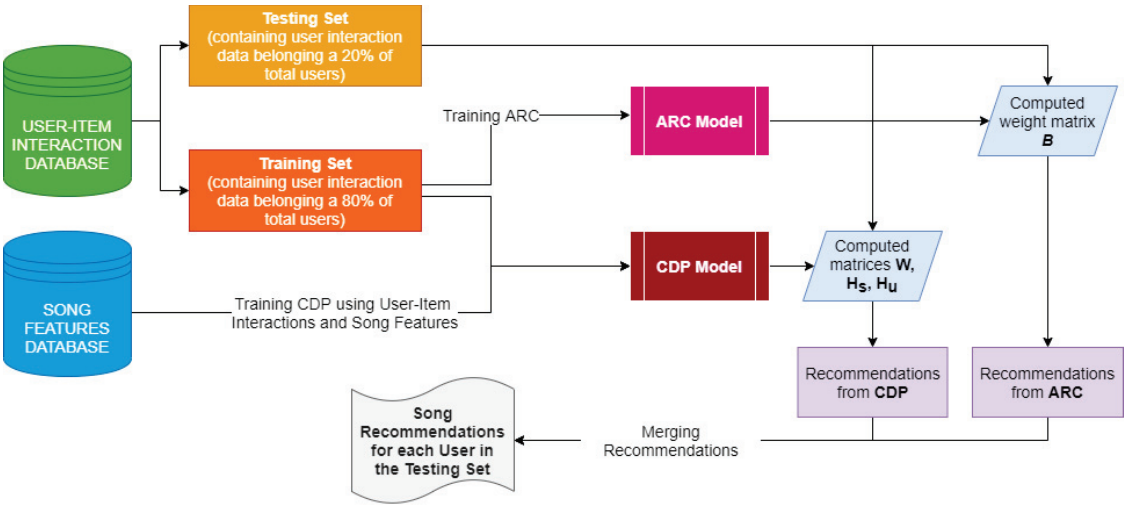


Fig. 2. Model Architecture

$q_u \in \mathbb{R}^{l \times u}$ which indicates how likely a user is expected to show interest in the item q.

As given in Saveski et. al[9], each item that is present in the dataset is associated with its description and a set of users who consumed them. In case of music data the description of the users can be considered equivalent to the metadata associated with each song. Hence, the input data is described previously as X_s and X_u. The matrix X_u is indicative of whether a user has interacted with an item or not similar to the representation in Equation 1. On the other hand, X_s is indicative of the inherent properties of each song and contain no information regarding their interaction with the users in the system.

Since both the matrices define independent latent spaces (since they convey different information), if treated in different expressions will show no correspondence whatsoever. The idea of *Local Collective Embedding* is that both item description and user interactions should be represented in a common latent space. So, there is a need to represent each factor in a expression (in this case the loss function for training given by J) that combines both the latent spaces. In our case, we achieve the common latent space by using the same lower dimension matrix W.

Once we define the training for CDP, we obtain the output vector of users for each item, hence solving the item cold start problem.

B. Model Training

In this section we shall take a look at the training procedure for each of the sub models of our recommender system that is required to be taken up in order to effectively generate recommendations of existing as well as new items or songs.

*1) **ARC**:* : We use the following objective function for learning the weights B:

$$\min_{B} \quad \|X - XB\|_F^2 + \lambda \cdot \|B\|_F^2 \tag{4}$$

$$\text{s.t.} \quad diag(B) = 0 \tag{5}$$

A few points to note here to understand the rationale behind choosing such a training procedure

- We choose the square loss function between the input data X and the predicted score matrix S as given in Equation 2 because it enables us to apply the closed-form solution easily in the upcoming steps.
- The L2-norm regularization has been used with the loss function. Hence the single hyperparameter for the sub-model is λ, that we aim to choose in an optimal manner.
- The constraint of removing self similarity by enforcing $\text{diag}(B) = 0$ is pertinent. If the constraint is ignored, we may obtain trivial solutions for $B(= I_n)$ where I_n denotes the identity matrix of order n

*a) **Deriving the Closed-Form Solution**:* The constrained optimisation problem given in equations 4 and 5 can be solved in closed-form, using a mathematical expression that uses a finite number of standard operations.

We first combine the target function and the constraints into a single optimisation function by forming the Lagrangian

$$L = \|X - XB\|_F^2 + \lambda \cdot \|B\|_F^2 + 2\gamma^T \text{diag}(B) \tag{6}$$

where $\gamma^T = \begin{bmatrix} \gamma_1 & \gamma_2 & \dots & \gamma_{|I|} \end{bmatrix}$ is the matrix of Langrange mulitpliers.

We shall choose the values for the elements of γ in a way that our equation 4 is optimised with the constraint in equation 5 being fulfilled.

The optimisation problem is solved by differentiating Equation 6 with respect to B and then setting the resultant equation equal to 0. Mathematically,

$$\frac{\partial L}{\partial B} = 0 \tag{7}$$

We know that for $[x_{ij}]\ \forall x_{ij} \in \mathbb{R}$, the following relationship holds true

$$\|X\|_F^2 = \text{Tr}(X \cdot X^T) \tag{8}$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:18:07 UTC from IEEE Xplore. Restrictions apply.

where $\text{Tr}(A)$ represents the Trace of matrix A. Using 8, we rewrite Equation 6, and the equation hence obtained after simplification is as follows:

$$\begin{aligned} L =& \text{Tr}(XX^T - XB^TX^T - XBX^T + XBB^TX^T) \\ &+ \lambda \text{Tr}(BB^T) + 2\gamma^T \text{diag}(B) \end{aligned} \tag{9}$$

When we differentiate Equation 9 with respect to B, we obtain the following expression:

$$\frac{\partial L}{\partial B} = -2X^TX + 2X^TBX + 2\lambda B + 2 \cdot \text{diagMat}(\gamma) \tag{10}$$

where diagMat(.) denotes the diagonal matrix. Setting value of $\frac{\partial L}{\partial B} = 0$ in Equation 10, we obtain the approximate value for B that we denote by $\hat{B}$, the value for which is represented as follows:

$$\hat{B} = (X^TX + \lambda I)^{-1} \cdot (X^TX - \text{diagMat}(\gamma)) \tag{11}$$

where I denotes the identity matrix. In order to simplify further calculations, we make the following substitution,

$$\hat{P} = (X^TX + \lambda I)^{-1} \tag{12}$$

Using this value of $\hat{P}$ in Equation 11, we get the following expression

$$\begin{aligned} \hat{B} &= \hat{P} \cdot (P^{-1} - \lambda I - \text{diagMat}(\gamma)) \\ &= I - \hat{P}(\lambda I + \text{diagMat}(\gamma))) \\ &= I - \hat{P} \cdot \text{diagMat}(\tilde{\gamma}) \end{aligned} \tag{13}$$

where $\tilde{\gamma} \triangleq \lambda \vec{1} + \gamma$

We now apply the Constraint 5 to the above value of B and obtain the following expression

$$0 = \text{diag}(\hat{B}) = \vec{1} - \text{diag}(\hat{P}) \odot \tilde{\gamma} \tag{14}$$

where $\odot$ denotes the elementwise product and hence derive the relationship

$$\tilde{\gamma} = \vec{1} \oslash \text{diag}(\hat{P}) \tag{15}$$

where $\oslash$ denotes the elementwise division (which is always defined since $\hat{P}$ is invertible). Hence we obtain the value for the Langrangian multipliers, and substitute the same value in Equation 11 resulting in the closed-form solution as follows:

$$\hat{B} = 1 - \hat{P} \cdot \text{diagMat}(\vec{1} \oslash \text{diag}(\hat{P})) \tag{16}$$

The above equation proves that the weights can be learned in closed-form without need for gradient analysis algorithms like Stochastic Gradient Descent.

The elementwise representation for $\hat{B}$ can be derived from 16 and is defined as follows:

$$\hat{B}_{i,j} = \begin{cases} 0 & \text{if } i = j \\ -\frac{\hat{P}_{ij}}{\hat{P}_{jj}} & \text{otherwise} \end{cases} \tag{17}$$

The values hence obtained for $\hat{B}$ obeys the Constraint given in Equation 5. We infer from Equation 17 that the non-diagonal elements for $\hat{B}$ can be calculated using the matrix $\hat{P}$ whose value is given in Equation 12.

Algorithm 1: Training the ARC Model

Input: User-Item Matrix := $X \in \mathbb{R}^{|U| \times |I|}$, L2-norm regularization-parameter $\lambda \in \mathbb{R}^+$

Output: weight matrix B with $\text{diag}(B) = 0$

```
G ← X^T X  /* dimension of G is |I| × |I|  */
for i ← 1 to |I| do
    G_{i,i} ← G_{i,i} + λ
P ← G^{-1}
for i ← 1 to |I| do
    for j ← 1 to |I| do
        B_{i,j} ← -P_{i,j} / P_{i,i}
for i ← 1 to |I| do
    B_{i,i} ← 0
```

Algorithm 1 explains how the training process runs on the given user interaction matrix using the formula derived for training B from Equation 17.

2) CDP: For CDP, we use the following optimization function, in accordance with our aim of obtaining a common latent space representation.

$$\begin{aligned} \min_{W,H_s,H_u} \quad J =& \frac{1}{2}(\alpha \|X_s - WH_s\|_F^2 \\ &+ (1-\alpha)\|X_u - WH_u\|_F^2) \\ &+ \lambda(\|W\|_F^2 + \|H_s\|_F^2 + \|H_u\|_F^2) \end{aligned} \tag{18}$$

such that $W \geq 0, H_s \geq 0, H_u \geq 0$

*a) **Localised Collective Factorization**:* The feature space that we work on exists in a higher dimension than at which we would want to analyse. To develop linear approximations for our matrix calculations and optimisations, it would be helpful if we could reduce dimensionality.

To do the same, we make use of the concept of *manifold assumptions*, similar to the one used in Luo et al[10]. The manifold assumptions is that, instead of assuming that data in the world can come from every part of the possible vector spaces, it is sensible to assume for training data to come from relatively low dimensional manifolds. If two distinct data points (vector of data features in this case) x_i and x_j, are close to each other in the higher dimension space, then the assumption that their representation in lower dimensions also lies close to each other is tolerable.

In order to reduce the dimensionality of our data, we perform the following steps as explained in Saveski et al and other literature [9], [11]:

1) Consider a graph G with n nodes where each node represents a vector in the feature space
2) For each point find the k Nearest Neighbours and connect corresponding nodes in the graph. The edges assume a value of 1 if they are among the k Nearest Neighbours and a value of 0 otherwise.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:18:07 UTC from IEEE Xplore. Restrictions apply.

3) This gives a matrix A as a representation of the cumulative graph which can be used to measure the local closeness (closeness in lower dimension) of data points

After reducing dimensionality, we form an expression in terms of the local closeness of points by taking the Euclidean distance between each row of the lower dimensional matrix **W** and also considering the matrix A.

This gives the equation:

$$S = \frac{1}{2}\sum_{i,j=1}^{n} \|w_i - w_j\|_F^2 A_{ij}$$

From [9], we obtain a reduced value for the above equation, given by:

$$S = \text{Tr}(W^T DW) - \text{Tr}(W^T AW) = \text{Tr}(W^T LW) \tag{19}$$

where D is a diagonal matrix whose entries are the row sums of A, i.e., $D_{ii} = \sum_i A_{ij}$; $L = D - A$ is called the Laplacian matrix of the graph[12] and Tr$(\cdot)$ is the trace of a matrix.

*b) **Objective Function**:* Hence the final objective function that we obtain is as follows:

$$\begin{aligned} \min_{W,H_s,H_u} \quad J = &\frac{1}{2}(\alpha\|X_s - WH_s\|_F^2 \\ &+ (1-\alpha)\|X_u - WH_u\|_F^2) \\ &+ \beta\text{Tr}(W^T LW) \\ &+ \lambda(\|W\|_F^2 + \|H_s\|_F^2 + \|H_u\|_F^2) \end{aligned} \tag{20}$$

such that $W \geq 0, H_s \geq 0, H_u \geq 0$; where L is the Laplacian matrix of the graph and β is a hyper-parameter which controls the extent to which locality has to be enforced.

*c) **Learning Procedure**:* Equation 20 is non-convex in nature when we take W, H_s, H_u together, a gradient descent algorithm can not be applied to reduce the objective function. We make use of a mathematical result as given in Cichocki et al[13], which states that in order to obtain a minimum stationary point, the following update rules should be applied to the matrices given a non convex objective function:

$$W \leftarrow W \odot \frac{\alpha X_s H_s^T + (1-\alpha)X_u H_u^T + \beta AW}{\alpha W H_s H_s^T + (1-\alpha)W H_u H_u^T + \beta DW + \lambda W} \tag{21}$$

$$H_s \leftarrow H_s \odot \frac{\alpha W^T X_s}{\alpha W^T W H_s + \lambda H_s} \tag{22}$$

$$H_u \leftarrow H_u \odot \frac{(1-\alpha)W^T X_u}{(1-\alpha)W^T W H_u + \lambda H_u} \tag{23}$$

where $\odot$ denotes element-wise multiplication of matrices.

C. Generating Recommendations

We train ARC and CDP using Algorithm 1 and 2. Once training has been completed, we obtain trained matrices B, W, H_s and H_u where the latter three are from the CDP model.

Post this, as given in Figure 2 we generate recommendations for the users on the testing set and evaluate our model as follows:

Algorithm 2: Training the CDP Model

Input: User-Item Matrix := $X_u \in \mathbb{R}^{n\times u}$, Item-Metadata Matrix := $X_s \in \mathbb{R}^{n\times m}$, L2-norm regularization-parameter $\lambda \in \mathbb{R}^+$, **hyper-parameter** $\alpha \in \mathbb{R}$ that controls the importance of factorization of X_s and X_u, **hyper-parameter** $\beta \in \mathbb{R}$ that controls the extent to which locality is enforced, number of iterations num_iter, k the number of top scores to be considered

Output: Trained matrices W, H_s, H_u

$W :=$ randomly initialised matrix $\in \mathbb{R}^{n\times k}$

$H_s :=$ randomly initialised matrix $\in \mathbb{R}^{k\times m}$

$H_u :=$ randomly initialised matrix $\in \mathbb{R}^{k\times u}$

$\gamma := 1 - \alpha$

for $i \leftarrow 1$ *to* num_iter **do**

/* As derived in Equations 21, 22, 23 */

$H_s \leftarrow H_s \odot \frac{\alpha W^T X_s}{\alpha W^T W H_s + \lambda H_s}$

$H_u \leftarrow H_u \odot \frac{(1-\alpha)W^T X_u}{(1-\alpha)W^T W H_u + \lambda H_u}$

$W \leftarrow W \odot \frac{\alpha X_s H_s^T + (1-\alpha)X_u H_u^T + \beta AW}{\alpha W H_s H_s^T + (1-\alpha)W H_u H_u^T + \beta DW + \lambda W}$

return W, H_s, H_u

- Each user u_i in the testing set has a corresponding vector consisting of past interactions with songs.
- First, we serve this vector as input to the ARC model, and obtain in the output layer S the scores for each item corresponding to the user as given in Equation 3. We select the top k scores in S and hence obtain the recommendations from the ARC.
- Second, we use the trained matrices H_s and H_u to get the score as follows.

$$W_{test} = X_s \odot H_s^{-1} \tag{24}$$

$$L = W_{test} \odot H_u \tag{25}$$

where W_{test} denotes the weight matrix relating song features with the lower dimensional representation of song features, and L denotes the scores for the users interacting with the items.

- Finally, we combine the top recommendations from both the ARC and CDP models to obtain the final k items.

Observe that the generation of item recommendations from the CDP rely on the song features and hence are able to specifically address the Item Cold Start Problem. We refer to Algorithm 3 to understand the procedure for generating the top k recommendations.

IV. Results and Analysis

In this section, we explore the results of the EASE'em model and compare its running performance with similar recommendation systems in literature that have made their results publicly available.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:18:07 UTC from IEEE Xplore. Restrictions apply.

Algorithm 3: Generating Recommendations from Model

Input: User-Item Matrix $X_u \in \mathbb{R}^{|U|\times|I|}$, Trained Weight Matrix for ARC $B \in \mathbb{R}^{|I|\times|I|}$, k the number of item recommendations to generate per user, Item-Features Matrix $X_s \in \mathbb{R}^{|I|\times m}$ where m is the number of features, Trained lower-dimensional representation for **CDP** W, H_s, H_u

Output: Matrix $Q \in \mathbb{R}^{|U|\times k}$: Generated recommendation for each user u

$S = X_u \cdot B$
$W_{test} = X_s \odot H_s^{-1}$ /* To solve for W in $X_s = WH_s$ */
for $w_{ij} \in W_{test}$ **do**
  if $w_{ij} < 0$ **then**
    $w_{ij} \leftarrow 0$ /* Set a threshold of 0 */
  else
    continue
$L := W_{test} \odot H_u$
$Q :=$ Null Matrix of dimension $|U| \times k$
for $u \leftarrow 1$ *to* $|U|$ **do**
  $row_L \leftarrow \text{sort}(L_{u,.})$ /* Sort row corresponding to user u in Q */
  $row_L \leftarrow row_L[-k:]$ /* Choosing the highest k elements */
  $row_S \leftarrow sort(S_{u,.})$ /* Sort row corresponding to user u in Q */
  $row_S \leftarrow row_S[-k:]$ /* Choosing the highest k elements */
  $row \leftarrow row_L \cup row_S$
  $Q[u] \leftarrow sort(row)$
return Q

As mentioned before, we apply our model on the Million Song Dataset[8].The overall description of the data used is as follows:

- Contains **76353** users and **10000** songs in the dataset
- Contains the information of interaction between each user and song. A value of 1 in the dataset indicates that the user has interacted with the item, whereas a value of 0 indicates non-interaction.
- Contains metadata information (**artist familiarity, artist hotness, sample rate, loudness, tempo, time signature, key, confidences etc**) of every song

The evaluation is conducted by splitting the data sets containing training and testing data respectively. The validation is conducted on the basis of *weak generalisation* where the training and testing sets do not have any common user-item interaction pairs. Then the training and testing data flow into the model as represented in Figure 2.

The evaluation metrics for model performance are as follows:

TABLE I
PERFORMANCE METRICS FOR EASE'EM

	K		
	20	**30**	**50**
Precision@K	0.216	0.197	0.14
Recall@K	0.4508	0.4941	0.5836
NDCG@K	0.401	0.365	0.322

- Precision@K, $P@K$: Refers to the fraction of correctly recommended items for every song recommended by the model. This metric has not been used by any other model, since it can be misleading metric to adjudicate a model on.
- Mean Average Precision@K, $MAP@K$: Mean Average Precision refers to the mean of all average precision values over the testing dataset. The expressions for average precision ($AP@K$) and mean average precision@K ($MAP@K$) is as follows:

$$AP@K = \frac{1}{r}\sum_{i=1}^{K} P@K \cdot rel_k \tag{26}$$

where rel_k checks if the kth item has been in the users actual interaction history and r refers to the number of such relevant items.

$$MAP@K = \frac{1}{|U|}\sum_{u=1}^{|U|} AP@K_u \tag{27}$$

where $|U|$ denotes the total number of users and $AP@K_u$ refers to the average precision for the uth user. $MAP@K$ is often used to evaluate the change in performance with the changing hyperparameters. *In our approach since the hyperparameter λ (in the ARC model) is influential, we observe in Figure 3 that as lambda increase the $MAP@K$ decreases.*
- Recall@K, $R@K$: The recall for a recommender refers to the fraction of correctly recommended items to the total number of songs heard by users.
- $NDCG@K$: The Cumulative Gain for a list of p items ranked from 1 to p, where the relevance of each item is denoted by rel_i is given as

$$CG_p = \sum_{i=1}^{p} \text{rel}_i \tag{28}$$

The cumulative gain does not consider the ranking of the items to be of any importance. To overcome this we use Discounted Cumulative Gain which is given as:

$$DCG_p = \sum_{i=1}^{p} \frac{rel_i}{\log_2 i + 1} \tag{29}$$

In order to compare the Discounted Cumulative Gain, we need a reference score. This is obtained by an ideal case for our recommendation model, which we refer to as the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:18:07 UTC from IEEE Xplore. Restrictions apply.

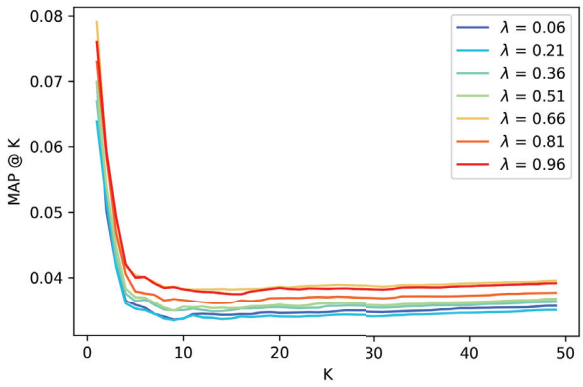


Fig. 3. Variation of MAP@K with λ

TABLE II
COMPARISON OF PERFORMANCE METRICS WITH NEURAL NETWORK BASED MODELS

	Recall@20	Recall@30	Recall@50	NDCG@50
EASE'em	**0.4508**	**0.4941**	**0.5836**	0.322
EASE [14]	0.33	0.394	0.428	**0.389**
Multi-VAE [15]	0.266	0.283	0.364	0.316
Multi-DAE [16]	0.266	0.313	0.363	0.313
RecVAE [17]	0.276	0.359	0.374	0.326
RaCT [18]	0.268	0.297	0.364	0.319
LambdaNet [19]	0.259	0.329	0.355	0.308

Ideal Discounted Cumulative Gain, $iDCG$ computed as follows:

$$iDCG_p = \sum_{i=1}^{|REL_p|} \frac{rel_i}{\log_2 i + 1} \tag{30}$$

where, $|REL_p|$ represents the size of the ideal recommendation generation list.
Finally, we obtain the $NDCG$ by taking the ratio of the DCG of a list to the $iDCG$ corresponding to the same list, and obtain the formula as,

$$NDCG = \frac{DCG_p}{iDCG_p} \tag{31}$$

As observed in Table I and II, the recall and NDCG of our model is far higher than the alternatives, which proves how our model effectively recommends good songs to the users, on an overall basis. This can be explained by our model's ability to effectively recommend songs with little or no previous user interactions, hence addressing the Item Cold Start Problem. We further notice that our precision values decrease with k and the recall values increase with k which is in coherence with the general understanding of the metrics.

First comparing our model with EASE[14]. By usage of autoencoders it suffers from the flaw of **unable to recommend songs that have no past user interactions** and also suffers from **Popularity Bias**. Further, song metadata is neglected. The NDCG for the EASE model is higher than our model, merely because **our model sacrifices giving more correct predictions by giving recommendations that are free (to a large extent) from popularity bias.**

Comparing with models that make use of autoencoders based on deep learning approaches and collaborative filtering namely Multi-VAE[15], Multi-DAE[16], RecVAE[17], we observe the following:

- The deep NN based approaches suffer from popularity bias because these models train on the same item that has a higher listen count.
- They constantly attempt to understand the entire complex relation between the features and user interactions of the songs with the output that is obtained.

The RaCT[18] model improves on the existing deep neural network and collaborative filtering, by amortising the item scoring process but it still attempts to interpret the complex relationship between the user-item matrix and item features matrix, which in turn increases the wall clock time for training without giving a better recall.

This is in contrast to our model, where we convert the feature space to a lower dimensions and then develop approximations for our model. When we observe that the linear approximations that we obtain after converting the feature space into a lower dimension give more relevant results, we can infer to a significant level the space where the features are much more simpler to interpret.

In summary we see that our model provides the following advantages over the other music recommender systems:

- Our model is able to effectively recommend songs that have little or no past user interactions. This means that an item with a large number of past user interactions do not necessarily skew the model into recommending that particular song to every other use. Hence, solving the *Item Cold Start Problem* and countering the *Popularity Bias*.
- We are able to achieve better overall performance with a lesser wall clock time for training.

V. CONCLUSION AND FUTURE WORK

Despite the relatively less amount of training time required for EASE'Em, our model was able to outperform various models on multiple parameters. On metrics at k=20 and k=50, EASE'Em stood as the best performing model, while on the NDCG metric it was only outperformed by the original EASE model. None of the previous papers concerned dealt with integrating cold start recommendations in their generations, and hence our model was able to uniquely tackle this problem. While evaluating model metrics on the full dataset might have some changes in the model outputs, metrics on a significant subset of the same have shown promising results. The integration of user based approaches with content based recommendation inherently faces the problem of data sparsity.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:18:07 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] Jonathan L Herlocker, Joseph A Konstan, Loren G Terveen, et al. "Evaluating collaborative filtering recommender systems". In: *ACM Transactions on Information Systems (TOIS)* 22.1 (2004), pp. 5–53.

[2] Ernesto G Birgin and José Mario Martınez. *Practical augmented Lagrangian methods for constrained optimization*. SIAM, 2014.

[3] Lei Hou, Xue Pan, and Kecheng Liu. "Balancing the popularity bias of object similarities for personalised recommendation". In: *The European Physical Journal B* 91.3 (2018), pp. 1–7.

[4] Markus Schedl, Hamed Zamani, Ching-Wei Chen, et al. "Current challenges and visions in music recommender systems research". In: *International Journal of Multimedia Information Retrieval* 7.2 (2018), pp. 95–116.

[5] Marcos Aurélio Domingues, Fabien Gouyon, Alıpio Mário Jorge, et al. "Combining usage and content in an online recommendation system for music in the long tail". In: *International Journal of Multimedia Information Retrieval* 2.1 (2013), pp. 3–13.

[6] P. Chiliguano and G. Fazekas. "Hybrid music recommender using content-based and social information". In: *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2016, pp. 2618–2622. DOI: 10.1109/ICASSP.2016.7472151.

[7] Xia Ning and George Karypis. "Slim: Sparse linear methods for top-n recommender systems". In: *2011 IEEE 11th International Conference on Data Mining*. IEEE. 2011, pp. 497–506.

[8] Thierry Bertin-Mahieux, Daniel P.W. Ellis, Brian Whitman, et al. "The Million Song Dataset". In: *Proceedings of the 12th International Conference on Music Information Retrieval (ISMIR 2011)*. 2011.

[9] Martin Saveski and Amin Mantrach. "Item cold-start recommendations: learning local collective embeddings". In: *Proceedings of the 8th ACM Conference on Recommender systems*. 2014, pp. 89–96.

[10] Fulin Luo, Yajuan Huang, Weiping Tu, et al. "Local manifold sparse model for image classification". In: *Neurocomputing* 382 (2020), pp. 162–173.

[11] Mikhail Belkin and Partha Niyogi. "Laplacian eigenmaps and spectral techniques for embedding and clustering". In: *Advances in neural information processing systems* 14 (2001), pp. 585–591.

[12] Fan RK Chung and Fan Chung Graham. *Spectral graph theory*. 92. American Mathematical Soc., 1997.

[13] Andrzej Cichocki, Rafal Zdunek, Anh Huy Phan, et al. *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation*. John Wiley & Sons, 2009.

[14] Harald Steck. "Embarrassingly shallow autoencoders for sparse data". In: *The World Wide Web Conference*. 2019, pp. 3251–3257.

[15] Diane Bouchacourt, Ryota Tomioka, and Sebastian Nowozin. "Multi-level variational autoencoder: Learning disentangled representations from grouped observations". In: *arXiv preprint arXiv:1705.08841* (2017).

[16] Dawen Liang, Rahul G Krishnan, Matthew D Hoffman, et al. "Variational autoencoders for collaborative filtering". In: *Proceedings of the 2018 World Wide Web Conference*. 2018, pp. 689–698.

[17] Ilya Shenbin, Anton Alekseev, Elena Tutubalina, et al. "RecVAE: A New Variational Autoencoder for Top-N Recommendations with Implicit Feedback". In: *Proceedings of the 13th International Conference on Web Search and Data Mining*. 2020, pp. 528–536.

[18] Sam Lobel, Chunyuan Li, Jianfeng Gao, et al. "Towards Amortized Ranking-Critical Training for Collaborative Filtering". In: *arXiv preprint arXiv:1906.04281* (2019).

[19] Christopher JC Burges. "From ranknet to lambdarank to lambdamart: An overview". In: *Learning* 11.23-581 (2010), p. 81.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:18:07 UTC from IEEE Xplore. Restrictions apply.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/357027867

Effect of Shear Walls on Tall Buildings with Different Corner Configuration Subjected to Wind Loads

Chapter · December 2021

DOI: 10.1007/978-981-16-6557-8_59

CITATIONS
0

READS
31

1 author:

Ritu Raj Ahirwar
Delhi Technological University
17 PUBLICATIONS **31** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Wind effects on tall buildings View project

All content following this page was uploaded by Ritu Raj Ahirwar on 17 December 2021.

The user has requested enhancement of the downloaded file.

Effect of Shear Walls on Tall Buildings with Different Corner Configuration Subjected to Wind Loads

Saransh Mahajan, Vikramaditya Yadav, Rahul Raj, and Ritu Raj

Abstract Many structures which are constructed at present times tend to be wind-sensitive because of their shape, slenderness, flexibility, and size. This leads to demand for more economical wind resistant designs of tall buildings. For the stability of high-rise structure, lateral forces like wind forces and earthquake forces play a critical role; however, in case of high-rise structure, wind loads are more critical than earthquake forces. Lateral load resisting systems like shear walls are used in construction to resist the effect of lateral forces on the building. This paper is focused on studying the effect of shear walls on the deflection of a tall building exposed to the wind load in lateral direction and observing the variance in the deflection that will occur on the three columns (windward, leeward, and central) of a high-rise building with and without using the shear walls in different scenarios using Bentley STAAD Pro software. Prototype buildings are assumed to be made up of RCC beams and columns that are designed on a software constitutes of 36 story (G + 35) having a height of 3.5 m of each floors, and response analysis of these prototype buildings has been done for different columns of building exposed to 0° wind incidence.

Keywords STAAD pro · Tall building · Wind loads · Square-shaped structure · Square-shaped structure with triangular corners · Square-shaped structure with recessed corners · Shear walls

1 Introduction

Since the ancient times, humans always admired tall structures although there is a growing need in present times due to exponential growth in the population for the construction of high-rise building and expansion in horizontal direction is not more a practical solution specially in the urban areas. Mainly, the construction of high-rise

S. Mahajan (✉) · V. Yadav · R. Raj · R. Raj
Delhi Technological University, New Delhi, India

R. Raj
e-mail: rituraj@dtu.ac.in

© Springer Nature Singapore Pte Ltd. 2022

A. K. Gupta et al. (eds.), *Advances in Construction Materials and Sustainable Environment*, Lecture Notes in Civil Engineering 196,
https://doi.org/10.1007/978-981-16-6557-8_59

Saving to your locker...

building is focused on four criteria mainly human comfort, stability, strength, and serviceability. So, there is a need for the use of shear walls in the tall building to provide resistance against lateral loads like wind and earthquake loads. RCC shear walls starts from the foundation and extends up to the roof level of the structure and are generally provided along the length and width of a building. In the designing of a tall structure, the use of shear walls is to provide strength and increase the stiffness of structure which will ultimately curb the damage occurring to the structure due to lateral loads by reducing the lateral sway of the building; shear walls are provided in a structure in symmetrical manner for distributing the inelastic deformation in a uniform way over the plan of a building rather than concentrating on particular walls.

Although the wind loads are estimated using the standard code of practice that are being used around the world, but the information provided is very minimal and given only for standard shapes like square plan or rectangular plan building exposed to wind. Various literature and researchers have come up with the analysis and effect of shear walls on the deflection due to wind loads subjected to a high-rise structure.

Various standard codes of practice for the design of building and estimation of design loads such as wind load, imposed load, and dead load are used across the world for a RCC building although these codes only provide minimal information for the design of high-rise structure that is only for standard and conventional shapes like square plan or rectangular plan, but these information are not exhaustive in nature [1–8]. Murali Krishna and Dr. Arunakanthi in 2014 used Etabs software to perform analysis for various different locations of shear walls in an unsymmetrical high-rise building in order to find the location which is effective, efficient, and ideal for shear walls by comparing results on various parameters like bending moment, torsion, and shear force [9]. Esmaili et al. in 2008 conducted study on one of tallest RCC building having 56 storys which was located in high seismic zone in which the shear wall system was used under lateral and gravity loads under irregular openings and the behavior of structural elements were recorded [10]. Amin and Ahuja conducted experiments on wind tunnel models of "L" shaped and "T" shaped building for calculating wind pressure distributions of different faces, and these experiments were conducted on two buildings of same plan area and height but having different cross-sectional shapes with wind direction varying from 0° to 180° at an interval of 150° [11]. Kwok, in 1988, conducted wind tunnel model tests in order to analyze the effect of shape of building on the wind-induced responses on a tall building having rectangular cross section which were performed in a long-wind and cross-wind directions with the help of chamfered, slotted, and horizontal corners [12]. Mukherjee et al., conducted a case study using k-ε method for N plan shape high-rise building for the determination of wind pressure coefficients for wind angle of incidence varying from 0° to 180° having an interval of 30° [13]. Sheng et al., investigate unsteady characteristics of a wind, and its complete information is like profile, intensity, fluctuation for the global, and local wind loads for a high-rise building structure using wind tunnel test at a scale of 1:300 for a tall building with a well-defined atmospheric layer [14]. Ahlawat and Ahuja, in 2015, performed experiments in order to investigate wind loads generated in interference and isolated condition, and the experiments were performed on models of tall buildings having "Y" plan

Saving to your locker...

shape in open-circuit wind tunnel [15]. Chakraborty et al., in the study carried out wind tunnel experiment for a "+" shaped tall building where the surfaces are exposed to wind angle for 0° and 45° and mean wind pressure coefficient was calculated using [16]. Nagar et al., in the paper studied the wind forces on unconventional plan shaped tall building although wind tunnel test also performed for "H" plan and square plan shaped tall building and for these models interference factors are calculated [17]. Yuan-Lung Lo et al., in 2019, performed experiments to find out the wind forced acting on two principal buildings with the aim to measure overturning moments in both cross-wind and along-wind by using dummy building located very closely to the principal building by using high-frequency balancer [18].

2 Methodology

The analysis of buildings is performed with the help of STAAD Pro, a general purpose civil engineering software which is widely used in order to model, design, and analyze various types of structures like bridges, buildings, towers, transportation, industrial, and utility structures. It helps to simplify and increase the efficiency of the engineering process.

2.1 Wind Flow

Design wind Pressure: The speed of wind increases with height and is approximately zero at ground level and increases to maximum in the atmospheric boundary layer at a height called the gradient height as given by Eq. 1. Above the mean ground level, the wind pressure can be calculated at any height in the structure by using the relationship between wind pressure and the wind speed.

$$P_z = 0.6 \times C_p \times V_z^2 \tag{1}$$

where P_z, V_z, and C_p are wind pressure (N/m^2), design wind speed (m/s) at a height z above the mean level, and pressure coefficient, respectively.

Design wind speed,

$$V_z = V_b * k_1 * k_2 * k_3 * k_4 \tag{2}$$

where V_b is the basic wind speed (m/s) given by Eq. 2, varies with zones of the country, and k_1, k_2, k_3, and k_4 are risk coefficient, terrain roughness and height factor, topography factor, and importance factor for the cyclonic region, respectively.

Force coefficient method: On any specific building, the wind load is calculated as per the formula given in Eq. 3.

Saving to your locker...

$$F = C_f \times A_e \times p_z \tag{3}$$

where C_f, A_e, and p_z are force coefficient for the building, effective area normal to wind direction, and design pressure of the wind, respectively.

Here, three types of prototype buildings are constructed which are having same base area as well as same height, but the cross section varies due to different corner configurations. In all the buildings, shear walls are applied in two ways, and their effects are analyzed.

2.2 Shear Walls

Shear walls are vertical elements which are constructed to resist in-plane forces mostly wind and seismic loads and are hence constructed in areas with high wind and seismic activity. Shear walls are large structures that are usually built-in high-rise structures that are large. For construction of shear walls, materials such as concrete and masonry are used.

Shear walls case-1 In this case of shear wall, the shear walls are applied on the corners of the building such that they cover the start and end portion of all the external faces of the building forming L-shapes at corners when seen in the plan view.

Shear walls case-2 In this case of shear wall, the shear walls are applied on the central columns of the building such that it forms a square when seen in the plan view.

2.3 Details of Prototype Buildings

2.3.1 Types of Buildings

1. Square-shaped structure (Fig. 1)
2. Square-shaped structure with triangular corners (Fig. 2)
3. Square-shaped structure with recessed corners (Fig. 3)

2.3.2 Building Parameters

- Height = 126 m
- No. of floors = 36 (G+35)
- Concrete grade = M-25
- Steel reinforce = Fe-415
- Basic wind speed (Delhi) = 47 m/s
- Live load considered = 2.0 kN/m^2

Saving to your locker...

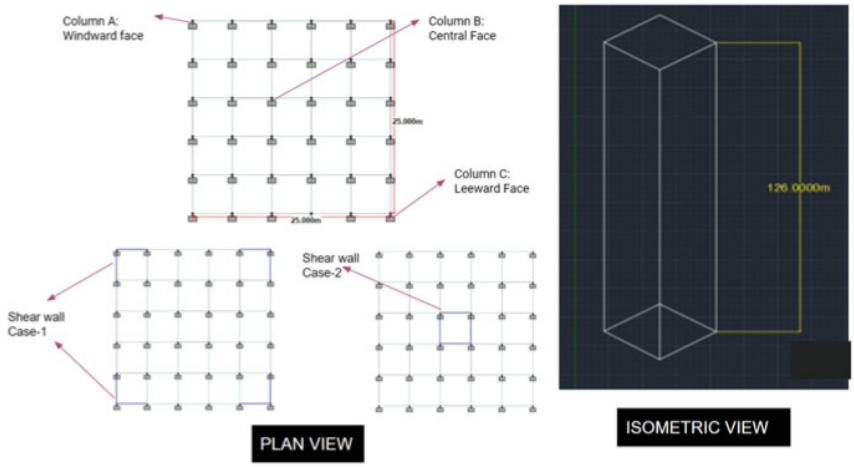


Fig. 1 Plan view of building, shear wall case-1, shear wall case-2, and isometric view of square-shaped building

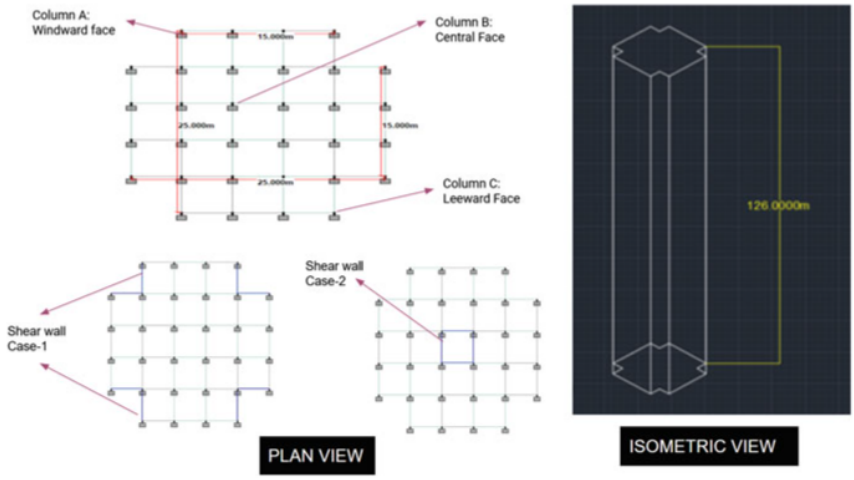


Fig. 2 Plan view of building, shear wall case-1, shear wall case-2, and isometric view of square-shaped building with triangular edges

- Floor finish = 1.0 kN/m^2
- Beam dimension = 450 × 450 mm
- Column dimension = 550 × 450 mm
- Main wall thickness = 230 mm
- Partition wall thickness = 150 mm
- Height of parapet wall = 1 m

Saving to your locker...

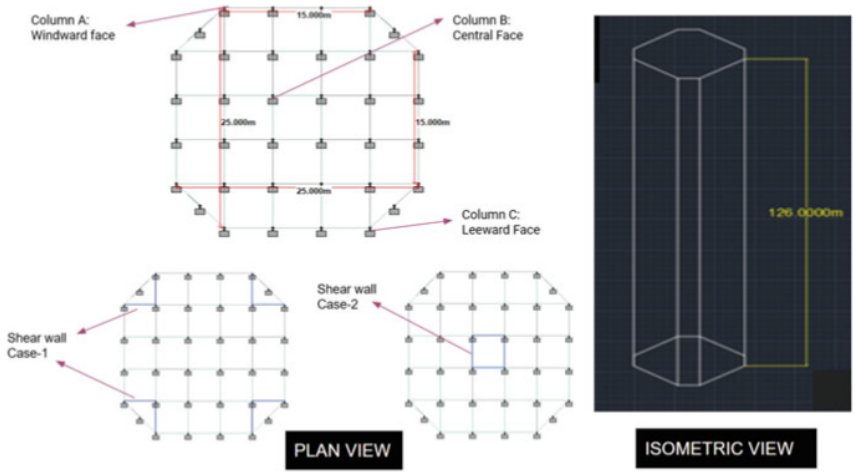


Fig. 3 Plan view of building, shear wall case-1, shear wall case-2, and isometric view of square-shaped building with recessed corners

- Slab thickness = 150 mm.

2.3.3 Loading Parameters (as per IS: 875 (Part I)—1987 [12], IS: 875 (Part II)—1987 [13])

- Self-weight = automatically generated
- Main wall load = 14.03 kN/m
- Partition wall load = 6.1 kN/m
- Parapet wall load = 4.6 kN/m
- Floor finish = 1 kN/m^2
- Floor live load = 2 kN/m^2.

3 Result and Discussion

3.1 Square-Shaped Building

The values of deflection increase as the height increases in all the three columns which are shown in Figs. 4, 5, and 6. When shear walls are applied, the values of deflection are found to reduce. Shear walls applied in first case gives better results than the shear walls applied in second case as they reduce deflection value by about 45%, while shear wall case-2 reduces the value by 34%. All the faces show similar

Saving to your locker...

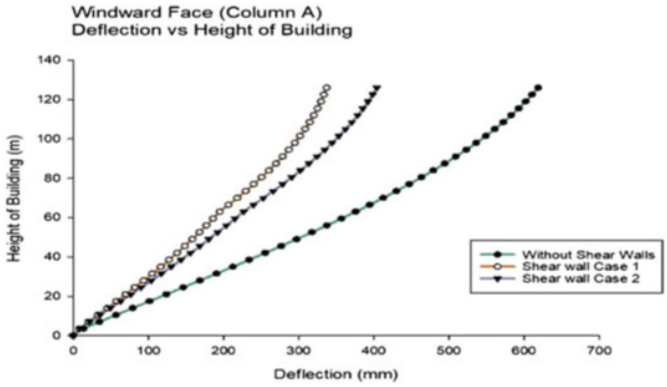


Fig. 4 Effect of shear walls on deflection at 0° incidence angle of wind on windward face (Column A) of square-shaped building

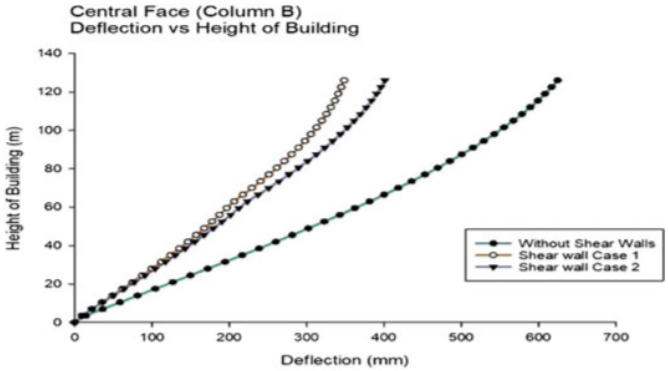


Fig. 5 Effect of shear walls on deflection at 0° incidence angle of wind on central face (Column B) of square-shaped building

values with slight difference in values with leeward face with maximum reduction first cases while central face in second case of shear walls.

3.2 Square-Shaped Building with Triangular Edges

The square-shaped building with triangular edges shows similar pattern (Figs. 7, 8, and 9) of increasing deflection with increasing height in all the three faces. Here, the shear walls in first case reduce the deflection by 21% with maximum in leeward face by slight margin, while shear walls in second case reduce deflection by 16% with slightly maximum reduction in central face.

Saving to your locker...

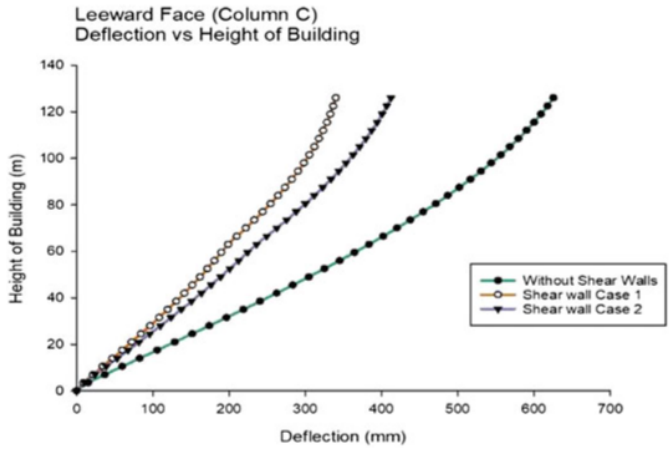


Fig. 6 Effect of shear walls on deflection at 0° incidence angle of wind on leeward face (Column C) of square-shaped building

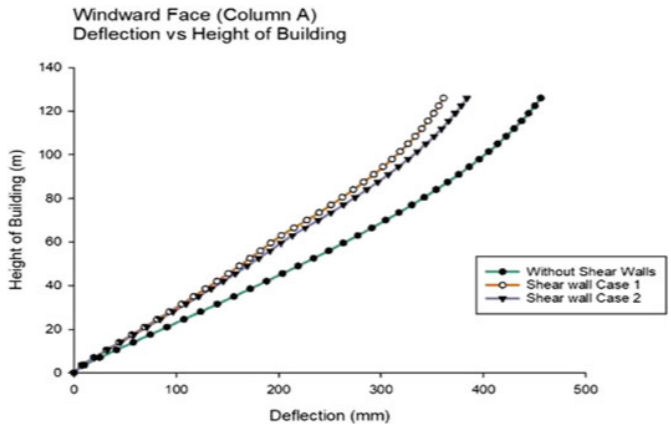


Fig. 7 Effect of shear walls on deflection at 0° incidence angle of wind on windward face (Column A) of square-shaped building with triangular edges

3.3 Square-Shaped Building with Recessed Corners

From Figs. 10, 11, and 12, it is found to increase in all three columns of square-shaped building with recessed corners with increasing height. The shear walls applied in first case helps to reduce the deflection by around 30% in leeward face and slightly less in other faces, while shear walls applied in second case helps to reduce the values of deflection by 18% in central face and slightly less in other faces.

Saving to your locker...

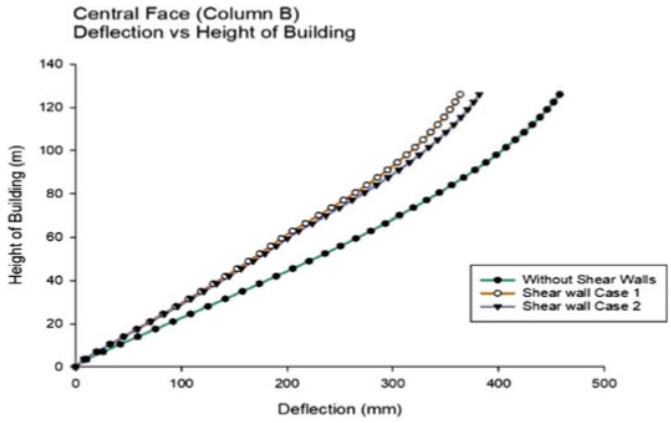


Fig. 8 Effect of shear walls on deflection at 0° incidence angle of wind on central face (Column B) of square-shaped building with triangular edges

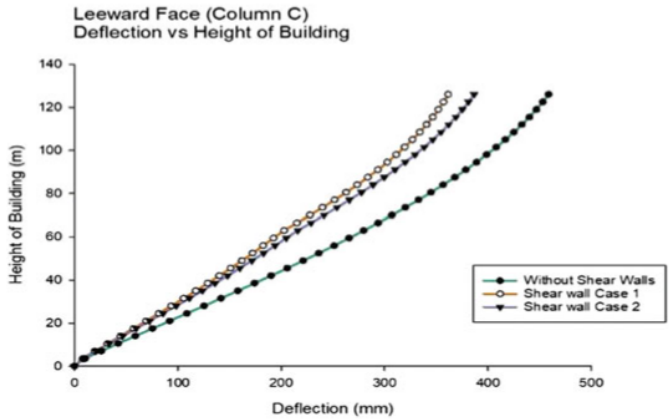


Fig. 9 Effect of shear walls on deflection at 0° incidence angle of wind on leeward face (Column C) of square-shaped building with triangular edges

4 Conclusion

1. The nature of graphs is found to be same in all three columns of all the buildings. The nature increases in deflection with increase in height of building.
2. All the three columns of any particular building show nearly same amount of reduction when shear walls of any one type are applied.
3. In square buildings, the first case of shear walls reduce deflection by about 45% while the second case of shear walls by about 34%.

Saving to your locker...

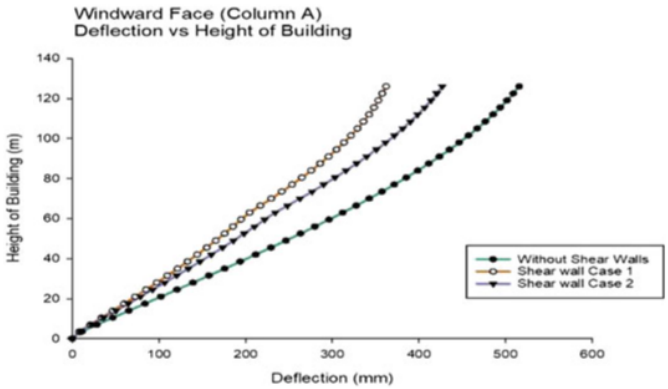


Fig. 10 Effect of shear walls on deflection at 0° incidence angle of wind on windward face (Column A) of square-shaped building with recessed corners

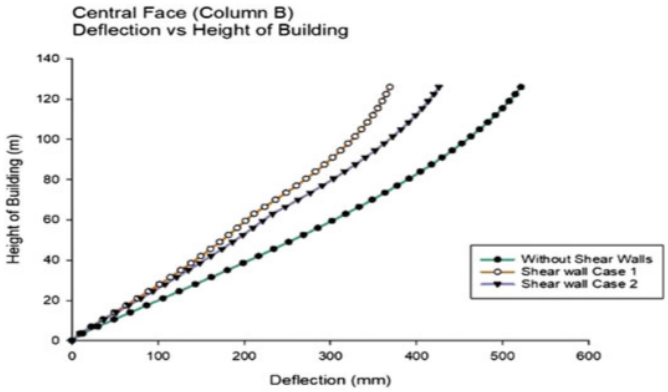


Fig. 11 Effect of shear walls on deflection at 0° incidence angle of wind on central face (Column B) of square-shaped building with recessed corners

4. In square buildings with triangular edges, the first case of shear walls reduces deflection by about 21% while the second case of shear walls by about 16%.
5. In square buildings with recessed corners, the first case of shear walls reduces deflection by about 30% while the second case of shear walls by about 18%.
6. The maximum reduction in deflection is seen in square building in both cases of shear walls in comparison with other two buildings.
7. Shear walls applied in first case (on the corners) are more effective in reducing deflection than the shear walls applied in second case (central columns) in all three buildings.

Saving to your locker...

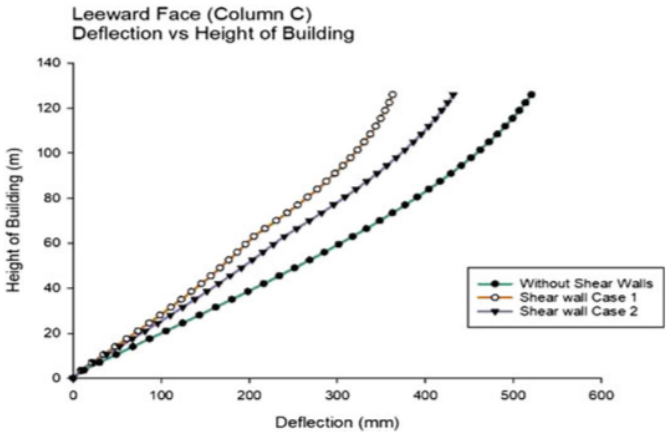


Fig. 12 Effect of shear walls on deflection at 0° incidence angle of wind on leeward face (Column C) of square-shaped building with recessed corners

Acknowledgements The authors gratefully acknowledge the technical support provided by the esteemed Delhi Technological University.

References

1. AS/NZS: 1170.2. (2002). Structural design actions, Part-2: Wind action.
2. ASCE: 7-02. (2002). Minimum design loads for buildings and other structures.
3. BS: 63699. (1995). Loading for buildings: Part 2-code of practice for wind loads.
4. EN: 1991-1-4. (2005). Euro code 1: Actions on structures wind actions.
5. IS:875-Part-3. (1987). Code of practice for design loads (other than earthquake loads) for buildings and structures-Wind loads.
6. IS:456. (2000). Plain and Reinforced Concrete—Code of Practice.
7. IS:875-Part-1. (1987). Code of Practice for Design Loads (other than Earthquake Loads) for Buildings and Structures—Dead Loads.
8. IS:875-Part-2. (1987). Code of Practice for Design Loads (other than Earthquake Loads) for Buildings and Structures—Imposed Load.
9. Krishna, M., & Arunakanthi, D. E. (2014). optimum location of different shapes of shear walls in unsymmetrical high rise buildings. *International Journal of Engineering Research & Technology (IJERT), 3*(9).
10. Esmaili, O. E. (2014). Study of structural RC shear wall system in a 56-Story RC tall building. In *14th World Conference on Earthquake Engineering.*
11. Amin, J. A., & Ahuja, A. K. (2008). Experimental study of wind pressures on irregular plan shape buildings. In *BBAA VI International Colloquium on: Bluff Bodies Aerodynamics and Application* (pp. 20–24).
12. Kwok, K. (1988). Effect of building shape on wind-induced response of tall building. *Journal of Wind Engineering and Industrial Aerodynamics, 28*, 381–390.
13. Mukherjee, A. B. (2017). Wind pressure and velocity pattern around 'N' plan shape tall building—A case study. *Asian Journal of Civil Engineering, 182*, 6399–6401.

Saving to your locker...

14. Sheng, R., Perret, L., Calmet, I., Demouge, F., & Guilhot, J. (2018). Wind tunnel study of wind effects on a high-rise building at a scale of 1:300. *Journal of Wind Engineering and Industrial Aerodynamics*, *174*, 391–403.
15. Ravinder Ahlawat, A. K. (2015). Wind loads on 'Y' plan shape tall building. *International Journal of Engineering and Applied Sciences (IJEAS), 2*(4).
16. Nagar, S. K., Raj, R., & Dev, N. (2014). Wind load on irregular plan shaped tall building—A case study. *Wind and Structures*, *19*, 59–73.
17. Nagar, S. K., Raj, R., & Dev, N. (2020). Experimental study of wind-induced pressures on tall buildings of different shapes. *Wind and Structures*, 31, 431–443.
18. Yuan-Lung Lo, Y. C. (2019). Estimation of wind-induced response of high-rise buildings immersed in interfered flow. *International Journal of Applied Science and Engineering, 22*, 429–448.

Check for updates

Research article

Exploring Linkages Between Workplace Spirituality and Occupational Stress Among Indian Police Personnel

South Asian Journal of Human Resources Management
1–26
© The Author(s) 2021
Reprints and permissions:
in.sagepub.com/journals-permissions-india
DOI: 10.1177/23220937211059626
journals.sagepub.com/home/hrm
SAGE

Stuti Jalan[1] and Naval Garg[2]

Abstract

The inclusion of spirituality at workplaces has yielded significant advantages for the employees and the organisations. The present study probes the optimistic facet of workplace spirituality (WPS) to identify its linkages with one of the most daunting issues for organisations, occupational stress (OS). Police officers from Himachal Pradesh, a hilly state in Northern India, were approached and responses were collected from 385 participants. The study utilises the three-dimensional WPS measure conceptualised by Milliman et al. (2003). Sense of community and value alignment was found to be explaining significant variance in occupational stress. The findings demonstrate the importance of embracing WPS as an effective tool for alleviating OS, thereby augmenting the performance and morale and guiding strategy formulation of future human resource development activities.

Keywords:

Workplace spirituality, occupational stress, stress, police

Submitted: 8 December 2020; **revised 1:** 13 June 2021; **revised 2:** 18 August 2021; **revised 3:** 8 September 2021; **accepted:** 16 October 2021

Introduction

Unprecedented changes in the work environment have mandated the decision-makers to reorient their thinking for improving performance, managing distress and increased dissatisfaction amongst the workers. Inclusion of a fourth

[1] Himachal Pradesh University, Shimla, Himachal Pradesh, India
[2] Delhi Technological University, New Delhi, India

Corresponding author:
Stuti Jalan, Himachal Pradesh University, Shimla, Himachal Pradesh 171005, India.
E-mail: stutijalan123@gmail.com

energy—'spiritual',—besides mental, physical and emotional energies of the employees has surfaced as a novel dimension to engage the whole self of an individual for better performance and help them to experience joy and improved engagement in their work (Dehler & Welsh, 2010). Owing to its consequential benefits to the individuals and the organisations, it has become a prominent theme of inquiry amongst scholars for identifying its exceptional role (Pawar, 2009). Empirical researches have supported its significant role in improving various performance indicators—job satisfaction (Garg et al., 2019; Pawar, 2009; Usman & Danish, 2010), organisational commitment (Garg, 2017; Rego & Cunha, 2008), work performance (Garcia-Zamor, 2003; Milliman et al., 2003), creativity and flexibility (Mitroff & Denton, 1999), innovation (Pandey et al., 2019), etc. These positive outcomes of WPS imply an improved work environment and employee satisfaction, which can also be presumed to be a significant indicator of lower occupational stress levels.

As far as WPS is concerned, it is defined as 'a framework of organisational values that are evidenced in the culture that promotes employees' experience of transcendence through the work process, facilitating their sense of being connected to others in a way that provides feelings of completeness and joy' (Giacalone & Jurkiewicz, 2003). Poe and Courter (1995) referred to WPS as 'the importance to people of finding deeper meaning in their lives (also their working lives) and the concomitant obligation of management to take care of their employees' needs in this regard'. In other words, WPS ensures alignment of body, mind and spirit at the workplace. WPS is a comprehensive term that takes care of the employees' physical, mental, emotional and spiritual needs (Dehler & Welsh, 2003). It is also noteworthy that spirituality and religion are not the same (Rathee & Rajain, 2020). Schneiders (2008) recommended three models to understand differences and linkages between spirituality and religion. The first model considers the two variables as separate and distinct enterprises. This model suggests no connection between the two. They believe that religion is more concerned with traditional values, rituals, customs and dogmatic beliefs. However, spirituality is more about aligning spirit (soul) with routine works. The second model articulates religion and spirituality as conflicting realities. The believer of the second model considers them in a reverse relationship, that is, more spiritual is less religious and vice versa. In this position, one may find religion hypocritic and start following spiritual practices for self-enlightenment. And finally, the third model suggests that religion and spirituality are two dimensions of a single enterprise. Such a model believes that both religion and spirituality are essential for one's awakening. They are complementary to each other.

Rapid changes and increased work pressure have made occupational stress pervasive in nearly all professions, including policing worldwide (Mathur, 1995; Nikam & Shaikh, 2014; Savery et al., 1993; Spielberger et al., 1981). A plethora of organisational studies present evidence on the negative implications of OS on physical (Pestonjee & Mishra, 1999; Sutherland & Cooper, 2000) and mental health (Cooper et al., 2001; Palmer et al., 2004) of individuals and on organisational outcomes. Job dissatisfaction, depression, alcoholism and even suicide are serious outcomes reported for occupational stress. The hidden costs associated with work

stress (increased medical expenditure, lower productivity) have required the managers and administrators to discover ways to handle, curb and control occupational stress.

For decades, individual spirituality has been considered an effective strategy to deal with personal and professional stress (Cox, 2011; Cunningham, 2014; Landis, 2009; Maheshwari, 2015; Mishra & Bisht, 2014; Powell et al., 2003; Zellars & Perrewe, 2003). However, the empirical studies examining the impact of spirituality in the workplace in managing occupational stress are sparse. Also, the studies identifying the linkages between WPS and OS have majorly examined academicians, students, nursing and other professionals (Bell et al., 2012; Chand & Koul, 2012; Csiernik & Adams, 2012; Daniel, 2015; Mahipalan & Suresh, 2019; Sharafi et al., 2014; Tischler et al., 2002). The present study explores the linkages between WPS and OS in a stress-prone occupation like policing, which is imperative for maintaining peace and order in any society. This research will offer valuable insights to the decision-makers on incorporating and embracing the spirituality dimension at the workplace to deal with one of the most critical issues faced by today's workforce—occupational stress (OS). Although previous studies, including Bell et al. (2012), Daniel (2015), Csiernik and Adams (2012), Cunningham (2014) and Sharafi et al. (2014), have explored the role of WPS in dealing with OS, the majority of studies are confined to Western settings. The present study aspires to validate the findings of these researchers in cross-cultural settings, that is, in the Eastern context.

Literature Review

Occupational Stress Among Indian Police Personnel

In India, policing is a highly sought job due to the power and prestige associated with it. Police personnel are treated with respect for their courage and sacrifices, which they display while maintaining law and order. But, they often experience high work stress due to insufficient funds and resources, lack of facilities, political interference, long and irregular working hours, inadequate staff, role conflicts and work pressure (Sharma, 1985, p. 64; Tyagi & Dhar, 2014). The policing job is also associated with high life risks, as police have to deal with dangerous criminals. The job profile is a psychologically draining job too. Because of the absence of a shift system, they may be called at night or during off-hours. Excessive work overload and job requirements deprive them of quality time for rejuvenation with family and friends, which exerts physical and mental strain on them. Paton and Violanti (1996) opine that police officers face higher burnout and lower morale than other professions. This work stress is manifested in the form of fatigue, anxiety, and depression. It negatively affects their behavioural responses and job satisfaction (Anderson et al., 2002).

Policing is treated as a 'critical occupation', which denotes that this profession is crucial for the well-being and protection of society. Presently, in the wake of COVID-19 and subsequent lockouts adopted by the nations to ensure social

distancing, their resilience has been further tested. Long working hours in a hostile work environment, the constant fear of infecting oneself and family members have compounded their existing levels of occupational stress. The emergencies and exhaustive routine functions of their work profile take a toll on their overall well-being resulting in occupational stress. The psychological well-being and mental health of officers should be a priority for the police department to help them handle distress caused due to various factors at the workplace (Khadse et al., 2020). It should be understood that the common motive in all their work activities is primarily aimed to maintain peace and harmony in society. Ironically, the stress faced by police officers prohibits them from having a peaceful state of mind, which can impact their performance and jeopardise the general public's safety.

The hilly terrain area of Himachal Pradesh puts extra pressure on the state police unit. Himachal Pradesh is a hilly state in northern India characterised by an extreme landscape. The state also receives snowfall and experiences low temperatures during winters, ranging between minus 8.2 degrees Celsius and minus 1.8 degrees Celsius in a few districts. Additionally, excessive rains result in landslides posing a challenge for the mobility of the general public. Hence, extreme weather conditions with limited resources coupled with rough terrain add to the woes of stressed police personnel. Moreover, it is a popular tourist destination, which witnesses millions of national and international tourists throughout the year. The state police must keep law and order to attract these tourists, who bring laurels to the state economy.

Police officers adopt different means to counter work-related stress. Some officers resort to maladaptive coping practices such as smoking, alcoholism, tobacco consumption, excessive eating and distancing from family to deal with extreme job stress (Almale et al., 2014). On the positive side, others opt for positive coping strategies such as listening to music, exercising, meditation, *pranayama* (breathing exercise), yoga and socialisation (Mathur, 1995). Although the police academies also organise seminars and workshops on work stress for police personnel, they are not accepted as routine activities. A few state police organisations experimented with spiritual training for their officers. Also, there is an acute scarcity of studies on the impact of WPS on OS among Indian police personnel. Police officers who received spiritual training experienced lower occupational stress (Pandya, 2017).

Despite the potential benefits of embracing workplace spirituality, the concept has failed to garner the required attention to deal with job stress in Indian police organisations. Accepting and nurturing the spiritual dimension of individuals and promoting spiritual values will create a healthy work environment for the employees, further contributing to developing personality attributes that help them to better cope with stress (Noor & Arif, 2011). Therefore, in light of the occupational stress in this critical occupation, it becomes even more imperative to explore the effect of WPS on OS to help organisations utilise the findings for improved morale and well-being of employees.

Workplace Spirituality and Occupational Stress

The idea of a spiritual workplace was conceptualised nearly a century ago in *Creative Experiences* by Mary Parker Follet (1930). Experiences at the workplace play a significant role in shaping an individual; hence, spiritual integration in the workplace activities during the eight hours could facilitate the spiritual development of an individual. Later, in 1977, Greenleaf suggested that spirituality has a significant role in the workplace (Neal, 2013, 5). Further, this thought was supported by Mitroff and Denton (1999), who argued that the human spirit is an essential element of one's existence, so leaving the spirit outside the work premises may be stressful for the employees (Pfeffer, 2003). It is crucial to integrate an employee's body, mind and spirit at work.

Meanwhile, rapidly changing economic trends have also accelerated the acceptance of spirituality at the workplace. Undesirable events like downsizing, layoffs, reengineering, longer working hours, mobility, etc., created a stressful environment. These stressed employees yearned for relief through a sense of connection and meaningful work, paving the way for a spiritual revolution in the workplace (Ashmos & Duchon, 2000). Interestingly, the factors that led to the spiritual revolution at the workplace were also considered major workplace stressors by both researchers and practitioners. For instance, earlier researchers have identified stressors like increased work pressure, long working hours, lack of transparency, instability, social isolation, etc. (Cartwright & Holmes, 2006; Colligan & Higgins, 2005). Later, these changes were argued as the drivers for the 'spirituality movement' at the workplaces (Ashmos & Duchon, 2000; Pfeffer, 2003). Therefore, the linkage between WPS and OS is evident in the conceptualisation of WPS itself.

Various factors intrinsic to the work environment (Cooper & Marshall, 1976) exhibit a poor fit between work demands and employee's abilities (Holmlund-Rytkönen & Strandvik, 2005). This employee–environment misfit results in occupational stress. Due to significant economic implications and adverse organisational and individual outcomes, occupational stress has become a severe concern for employers (Kelloway et al., 2008). Apart from severe health concerns, it can erode an individual's self-esteem, confidence, and capabilities (Maslach & Leiter, 1997). Besides lower productivity, these disruptions cause a change in individuals' behaviour, affecting the relationship with internal and external members associated with the organisation. It is imperative to explore various interventions to deal with the menace of OS. Expanding the basic premise, previous researchers have provided empirical evidence of a negative association between WPS and OS.

Karakas (2010) reported that workplace spirituality leads to increased morale, commitment, and decreased stress and burnout amongst employees. Zou and Dahling (2017) noted that spiritual inclination in an individual helped to preserve their well-being. Lately, researchers have investigated and found a positive influence of spiritual values/practices in the workplace on stress and mental well-being (Arnetz et al., 2013; Jnaneshwar & Sulphey, 2021; Kumar & Kumar, 2014; Saxena et al., 2020). Similarly, Goyal et al. (2014) also observed that spiritual

practices could ameliorate stress levels and improve work outcomes. Spiritually inclined employees perceived the organisational environment as healthy, safe, and pro-employee (Csiernik & Adams, 2012). Also, WPS is a more influential variable than job satisfaction and emotional ownership in dealing with job stress (Chand & Koul, 2012). The development of a spiritual angle in work processes, fostering interpersonal relationships with co-workers and focusing on greater social good would promote spiritual well-being and help handle stress (Lal et al., 2017; Saini, 2017). Researchers have proposed the three most commonly appearing dimensions in WPS literature—meaningful work, sense of community and alignment with organisational values (Ashmos & Duchon, 2000; Milliman et al., 2003; Rego & Cunha, 2008) form the framework of the present study. Henceforth, the author highlights the nexus identified by scholars between these spiritual dimensions and ideologies of the Indian context.

India has an ancient association with spirituality. Various spiritual concepts, rituals and practices are contributed by Indian *guru*s like Maharishi Ved Vyas, Kanad, Swami Vivekanand, Rajneesh Osho, Sri Ravi Shankar, etc. (Pandey et al., 2016; Sarkar & Garg, 2020; Saxena et al., 2020). Spiritual wisdom nourished by learned *guru*s and ancient scriptures of India provides a strong foundation for developing spiritually oriented management practices (Bhawuk, 2011). The Indian notion of *karma yoga* (selfless action performed for the benefit of others) gives centrality to work. Thus, the integration of spirituality at work is entirely natural to the Indian context. One of the most revered scriptures, Shrimad Bhagavad Gita, was preached on the battlefield of Kurukshetra. *Karma yoga* recommends a hundred per cent devotion to work without any feverishness of results. It states that one should not bother about results after the completion of the task. The Indian perspective inherently promotes a stress-free approach to work through encouraging work with the fear of failure. As discussed, the three most popular dimensions of WPS are meaningful work, a sense of community, and alignment with organisational values. It would be interesting to see the Indian perspective of these three modern constructs of WPS.

Meaningful Work

Finding meaning and purpose in one's work activities has been argued as a prerequisite for nurturing workplace spirituality (Milliman et al., 2003). Loss of meaning of work has left the employees dissatisfied in modern organisations (Bosch, 2009). Apart from several other factors like workload and shift duty, meaninglessness in the work activities contributed to job stress (Wu & Wong, 2008). Folkman (1997) observed that those engaged in meaning-based coping depicted a positive psychological state. Robertson and Cooper (2011) observed that employees who experienced a sense of accomplishment in their work displayed better mental state and psychological well-being. Engaging in meaningful activities helps 'control, manage, and prevent stress at the workplace' (Daniel, 2015).

The notion of 'meaningful work' has particular relevance in the Indian context. It is referred to as 'swadharama' (one's own duty) in the Indian tradition (Garg, 2017; Pandey et al., 2019). Various references to the word 'swadharama' could be found in multiple Indian scriptures like Shrimad Bhagwad Gita, Ramayana, Kena

Upanishad, Shiv Purana, etc. (Bhawuk, 2011). Also, the principle of Karma Yoga is regarded as the core aspect of Indian work culture. The idea of karma yoga also talks about the relevance of meaningful work. It entails duties to be performed with the most significant commitment to realise meaningfulness and purpose in work activities (Pardasani et al., 2014). It also suggests focussing on the process with dedication rather than outcomes.

A few Indian studies have explored the interaction between meaningful work and work stress. Pande & Naidu (1992) concluded that employees who focussed on efforts more than rewards experienced lower stress. Saxena et al. (2020) concluded a significant negative relationship between swadharama and stress among oil and gas exploration industry employees. They also realised that meaningful work is more effective in dealing with stress in the more challenging offshore situation. It is a noteworthy fact that offshore employees are more stressed than onshore employees. Like police personnel, offshore employees also work in unsafe conditions. There are a more significant number of accidents and causalities on offshore sites. They face extreme climatic conditions and psychological alienation (Kumar and Kumar, 2014). Drawing an analogy from their findings, we propose that WPS may counter OS in more challenging policing jobs. Hence, it was hypothesised that:

H_1: A higher level of involvement in meaningful work is related to lower occupational stress.

Sense of Community

Sense of community is 'a feeling that members have of belongingness, a feeling that members matter to each other as well as to the group, and a shared belief that members' needs will be sought after through their commitment to being together' (McMillan & Chavis, 1986). Being a part of the community and the urge for belongingness were identified as core elements for a spiritual revolution in the workplaces (Ashmos & Duchon, 2000). As human beings, individuals want to experience connection with others in their surroundings. With the increased centrality of workplaces, seeking relationships in the workplace has become more critical than ever. Eisler and Montouri (2002) discuss the growing focus of social relations in the organisations wherein the partnership model argues that the connections in the organisation improve attachment, including the strains, challenges and responsibility associated with the work.

The concept of fellowship and brotherliness in society has been of paramount importance since the Vedic times of India (Srirangarajan & Bhaskar, 2011). Maha Upanishad (verse 6.71–75) suggests '*tuvasudhaivakuṭumbakam*', which means the entire world is our family. This ancient teaching recommends the world as a single community having a universal consciousness. The presence of joint families, an extended web of relatives and stable marriages reflect a deep acceptance of the notion of a sense of community in the Indian context. Hindu tradition believes that all rituals and spiritual ceremonies elicit desirable fruits only if they are organised in the presence of all family members and friends. Collectivism and brotherhood are even more ingrained in *Pahari* culture (the

local culture of Himachal Pradesh). Their folks, rituals, festivals, customs and practices promote collective growth and development.

Modern researchers argue that the lack of social support and unpleasant relationships with fellow workers leads to alienation and a negative attitude towards work (Martinussen & Richardsen, 2006). Lazarus (1991) suggested removing isolation from the work environment to manage workplace stress. It is argued that the feeling of community is essential for seeking care, concern, and guidance at the workplace. Cooperation, coherence and social support can act as the fundamental principles for preventing OS in various workplace settings (Karasek et al., 1982; Michie, 2002) and promote to build a conducive work environment characterised by cooperativeness, teamwork, trust and strain reduction (Naylor et al.,1996). The findings discussed above are majorly derived from Western settings. It would be interesting to explore the association between WPS and OS in the Indian context, especially in the *Pahari* cultural context. Against this background, it was hypothesised that:

H_2: A higher level of sense of community is related to a lower level of occupational stress.

Alignment with Organisational Values

In light of the person–organisation fit theory, Kenny and Cooper (2003) argued that the discrepancy between organisational and individual values results in occupational stress. A better fit would result in increased satisfaction and lesser stress (French et al., 1982). Kolodinsky et al. (2008) argue that strong agreement between the individual and organisational values positively affects the employees' opinions, satisfaction and performance. A discrepancy in the value system depicts a misfit of the person and the organisation, leading to lower satisfaction levels. Kahn et al. (1964) describe that lack of consensus between individual beliefs and organisational values creates a stressful work environment. Also, Edwards (1992) reported discrepancies between personal and work values to be a significant source of stress.

Indian spiritual tradition has always supported synchronising one's value with that of organisation and society. 'It is better to do one's duty, though maybe devoid of merit, than doing the duty of others' (Shrimad Bhagwad Gita, Chapter 3, verse 35). It recommends the alignment of professional work in accordance with one's values system. It also advocates the distribution of organisational work according to one's inherent qualities and values (Pandey et al., 2009). Also, the ancient notion of '*daivi sampat*' (divine nature) seems to favour harmonising personal and organisational values. It encourages respect for co-workers and all other stakeholders of an organisation. It suggests that an ethical-driven work environment may lead to a positive organisational climate (Suresh & Janaki, 2012). Mahatma Gandhi (Father of the Nation, India) also stressed aligning work and personal values as a prerequisite for effective performance (Srirangarajan & Bhaskar, 2011). Alignment promotes increased belongingness wherein organisational goals become a priority over the personal purposes of individuals. The Indian value system has always preferred societal work over individual pursuit.

Management practices that promote integrated work life and do not conflict with an individuals' set of beliefs will help build a spiritual workplace (Pfeffer, 2003). Spiritual value-driven organisations can engage complete employees, involving their mind, body and soul (Pfeffer, 2003). Allowing integration of one's spiritual values in the workplace results in increased satisfaction and joy among employees; furthermore, the consensus between individual and organisational values yields better performance (Dehler & Welsh, 1994), and the disagreement between the value system results in anxiety culminating in stress and burnout which eventually affects their well-being (Maslach & Leiter, 1997). The proposed alignment of values is more critical for police personnel for two fundamental reasons. First, the inherent nature of a police job requires courage, determination, passion and sacrifice. A police officer is expected to align with organisational and societal expectations of being honest, transparent, impartial, patient, tolerant, approachable and polite. A corrupt or partial police officer is bound to be under stress with a constant fear of exposure or arrest. Second, police officers are entitled to various powers. And any misuse of such powers is detrimental for the entire society. It is an evident fact that stressed people easily engage in destructive activities. Any undesirable activity by police personnel hurts the officers and the reputation of the police force and society. Thus, alignment of organisational and personal values is highly desirable for police officers. Hence, it was hypothesised that:

H_3: Higher alignment with organisational values is related to lower levels of occupational stress.

Research Methodology

Sample and Data Collection

The present study is based on a cross-sectional research design, and data was collected with the help of stratified sampling. The population of the study was non-gazetted police officers of Himachal Pradesh. Data was sought from five strata of officers, that is, Inspectors, Sub-Inspectors, Assistant Sub-Inspectors, Head Constables, and Constables. Also, four districts of the state were chosen randomly for data collection, that is, Kangra, Shimla, Mandi and Solan. Nearly 500 respondents were approached, out of which 395 agreed to respond. Ten responses were discarded for being incomplete. Thus, the actual sample size of the present study is 385. The demographic distribution of the sample is depicted in Table 1.

Research Instruments

Workplace spirituality was measured using the Spirituality at Work (SAW) scale developed by Milliman et al. (2003). The scale consists of 21 items measuring workplace spirituality in three dimensions: meaningful work, sense of community

Table 1. Demographic Distribution of Sample.

Demographic Variables	Categories	*N*	Percentage
Gender	Male	282	73.2
	Female	103	26.8
	Total	**385**	**100**
Marital status	Married	262	68.1
	Unmarried	123	31.9
	Total	**385**	**100**
Age	21–31	155	40.2
	32–42	113	29.3
	43–53	63	16.3
	54–64	54	14.0
	Total	**385**	**100**
Educational qualification	Postgraduate	61	15.8
	Graduate	200	51.9
	Intermediate	80	20.8
	Matriculation	44	11.4
	Total	**385**	**100**
Job experience	2–10	190	49.3
	11–19	113	29.3
	20–28	56	14.5
	29–37	26	6.8
	Total	**385**	**100**
Rank	Inspector	07	1.8
	Sub-inspector	16	4.2
	Assistant sub-inspector	32	8.3
	Head constable	62	16.1
	Constable	268	69.6
	Total	**385**	**100**
Grade	Grade I	55	14.2
	Grade II	330	85.8
	Total	**385**	**100**

Source: Primary data.

and alignment with organisational values. All the items in the SAW scale were measured on a five-point scale from 1 (strongly disagree) to 5 (strongly agree). This scale was chosen for three fundamental reasons. First, it accesses workplace spirituality on individual (meaningful work), group (sense of community) and organisational levels (alignment with organisational values). Since the profession of efficient policing is an interplay between individual, group and organisational elements, the scale is very well suited to guide future workplace spirituality initiatives for police personnel. Second, although the scale has its origin in the Western context, psychometric properties are well-established in the Indian context (Pandey et al., 2009; Saxena et al., 2020). And lastly the nexus between these three spiritual dimensions and Indian traditions have been reported by previous scholars (Pandey et al., 2009; Pardasani et al., 2014; Saxena et al., 2020; Srirangarajan & Bhaskar, 2011). Thus, the SAW scale is a well-fit scale to measure the workplace spirituality of the Indian police system. A few sample items are 'I see a connection between my work and social good', 'I believe employees genuinely care about each other' and 'I feel connected with the goals of the organisation'.

Occupational Stress

Occupational stress was measured using the Organisational Role Stress (ORS) scale developed by Pareek (1983). It has been considered a classic inventory and is widely used by researchers to measure occupational stress in various professions (Ratna et al., 2013). The scale measures role stress in ten dimensions: Inter Role distance, Role Stagnation, Role Expectation conflict, Role Erosion, Role Overload, Role Isolation, Personal Inadequacy, Self-role distance, Role Ambiguity and Resource Inadequacy. All fifty items were measured on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The scale well captures the individual and organisational aspects of work stress. Thus, the scale captures stressors associated with policing duties. For instance, the imbalance between work and family (Inter Role Distance), exhaustive list of responsibilities (Role Overload), inadequate staff and limited resources (Resource Inadequacy), triple accountability to the department, government and general public (Role Expectation Conflict) are some of the factors that undeniably cause job stress among police personnel.

Reliability, Validity, Multicollinearity, and Common Method Bias (CMV)

The reliability of the scales was accessed with the help of Cronbach's alpha and composite reliability values. Previous researchers recommended an acceptable value of 0.70 for both reliability estimates. Table 2 reports decent values of Cronbach's alpha (Cronbach's alpha value for meaningful work = 0.865, sense of community = 0.816, alignment with organisational values = 0.890, occupational stress = 0.940). Also, composite reliability estimates for meaningful work, sense of community, alignment with organisational values and

Table 2. Descriptive Statistics.

Variable	Mean	*SD*	CA	*CR*	*AVE*	VIF
Meaningful work	2.94	0.96	0.865	0.899	0.649	2.148
Sense of community	2.77	0.90	0.816	0.866	0.597	2.120
Alignment with organisational values	2.97	0.92	0.890	0.912	0.519	2.601
Overall workplace spirituality	2. 89	0.95	0.912	–	–	–
Occupational stress	3.93	0.63	0.940	0.948	0.565	1.000

Source: Primary data.

Note: *SD*: Standard deviation, CA: Cronbach's alpha, CR: Composite reliability, AVE: Average variance extracted, VIF: Variance inflation index.

occupational stress are 0.871, 0.818, 0.894 and 0.948. Further, convergent validity was assessed using Average Variance Extracted (AVE), whose acceptable value is above 0.50. AVE values of meaningful work, sense of community, alignment with organisational values and occupational stress are 0.649, 0.597, 0.519, and 0.565 respectively. These values confirm the convergent validity of both scales. Also, the factor structure of the WPS scale has been added in the Annexure.

The chances of multicollinearity may arise owing to high correlations among variables of the study. This issue is examined using Variance Inflation Factor (VIF) values. VIF values less than five indicate insignificant multicollinearity among variables (Kline, 2005). Table 2 reports that all VIF values are below five (meaningful work = 2.148, sense of community = 2.120, alignment with organisational values = 2.601). Also, there are chances of CMV because both independent and dependent variables were collected simultaneously. Both procedural and statistical remedies are used to minimise CMV. First, the present study is based on pre-validated scales of WS and OS. And according to Podsakoff et al. (2012), it reduces the impact of CMV. Second, respondents were assured of anonymity and confidentiality of the collected data. They were also promised that the data would not be shared with anyone and used for academic purposes only. Such an assurance of secrecy of information encouraged respondents to respond more freely and honestly, reducing CMV chances (Podsakoff et al., 2012). Third, a separate cover story is given to WPS and OS to create psychological separation in the mind of respondents. They were assured that WSP and OS were part of two different studies, and the variables would not be evaluated simultaneously. Podsakoff et al. (2012) highly recommended such psychological separations of the study variables to deal with CMV. However, after completing the study, they were informed of the truth and the reason behind the adoption of psychological separation. And lastly, Harman's one-factor is used, wherein all variables of the study are subjected to exploratory factor analysis using principal component analysis. More than one factor emerges out, and the variance explained accounted for the largest factor is 30.565 per cent, which is less than the recommended threshold of 50 per cent. It concludes that CMV is not statistically significant for the present study.

Statistical Analysis

The interrelationships between meaningful work, a sense of community and alignment with organisational values with occupational stress were examined with the help of correlation and regression analysis. Also, stress levels of police personnel belonging to different demographic variables were assessed using the *t*-test and ANOVA and the demographic variables, for which stress variations are statistically significant, are separately subjected to regression analysis.

Results

Table 2 illustrates the mean, standard deviation, reliability, and validity estimates of the study variables. The respondents report themselves high on occupational stress with a mean value of 3.93. In terms of workplace spirituality, the respondents found their workplace as moderately spiritual (mean = 2.89). The mean values of meaningful work, sense of community and alignment with organisational values are 2.94, 2.77 and 2.97, respectively. Table 3 represents Pearson's correlation that assesses the relationship between workplace spirituality and occupational stress. Occupational stress is inversely and significantly correlated with workplace spirituality, with an r-value of −0.621. The size of correlation indicates moderate to high relationships between meaningful work (r = 0.423, p = .01), sense of community (r = −0.588, p = .01); alignment with organisational values (r = −0.608, p = .01). One interesting finding is the significant positive association of meaningful work with OS. Tables 4 and 5 provide the result of hierarchical regression that considers three dimensions of WPS as independent variables. In hierarchical regression, all spiritual dimensions are added in Model 1, and after that, all demographic variables are included in Model 2. The result suggests that the regression coefficient of meaningful work is insignificant at a 0.05 level of confidence. It means meaningful work does not predict OS, and hence, H_1 is rejected. The other two dimensions of WPS, sense of community ($\beta = -0.366$, $p <$ .05) and alignment with organisational values (β = −0.407, $p <$ .05), have a significant negative regression coefficient. It means H_2 and H_3 are accepted. Also, it is reported that the three dimensions of WPS could explain 42.4 per cent variations in OS in Model 1. Further, Model 2 suggests that apart from a sense of community and alignment with organisational values, grade and experience are also significant predictors of OS. It is also reported that Model 2 explains 48.9 per cent variations in OS. Table 6 reports the results of demographic variable-based variation in stress experienced by police officers. It is observed that OS varies with age, experience, educational qualifications, rank and grade of police personnel. The results reveal that OS increases in the initial years of services, and then it starts decreasing. The mean values of OS for employees belonging to age categories of 21–31 years, 32–42 years, 43–53 years, and 54–64 years are 3.99, 4.09, 3.56 and 3.40. The results are quite natural, as employees need time to settle down. It is also interesting to note that OS increases with the rank of the police personnel. Also, Grade-I (mean = 4.00) personnel experience higher OS levels

than officers of Grade-II (mean = 3.34). Obviously, with promotion, responsibilities and duties manifold, which may lead to increased stress levels. It is also reported that employees with high educational qualifications experience higher levels of OS. Table 7 represents regression equations of WPS and OS for different categories of demographic variables. Here, only variables with significant variations in OS are considered. It is observed that WPS is more effective for more stressed employees. WPS could explain 49.2 per cent and 18.8 per cent variations in OS among Grade-I (stress mean = 4.00) and the Grade-II police officer (stress mean = 3.34). Similarly, coefficient of determinations have higher values for employees with higher academic credentials (for PG officer = 36.9 per cent, Graduate = 32.3 per cent, 12th = 31.8 per cent, and matriculate = 7.1 per cent).

Table 3. Correlation Matrix.

Variables	1	2	3	4	5
Meaningful work	1	0.625**	0.709**	0.864**	0.423**
Sense of community			0.705**	0.860**	–0.588**
Alignment with organisational values				0.929**	–0.608**
Overall workplace spirituality					–0.621**
Occupational stress					1

Source: Primary data.

Note: ** correlation is significant at the 0.01 level.

Table 4. Results of Hierarchal Regression.

Dimensions	Unstandardised Coefficient	Standard Error	Standardised Coefficient	*t*-value	Sig.
Model 1					
Constant	5.244	0.114		45.925	0.000
Meaningful work	0.060	0.044	0.091	1.363	0.174
Sense of community	–0.249	0.050	–0.366	–5.004	0.000*
Alignment with organisational values	–0.284	0.046	–0.407	–6.165	0.000*
Model 2					
Constant	4.522	0.377		11.989	0.000
Meaningful work	0.036	0.042	0.054	0.853	0.395

(Table 4 continued)

(Table 4 continued)

Dimensions	Unstandardised Coefficient	Standard Error	Standardised Coefficient	*t*-value	Sig.
Sense of community	–0.233	0.048	–0.341	–4.868	0.000*
Alignment with organisational values	–0.240	0.045	–0.343	–5.297	0.000*
Rank	0.042	0.066	0.053	0.639	0.524
Grade	0.381	0.148	0.192	2.578	0.010*
Education	–0.031	0.038	–0.041	–0.803	0.422
Age	0.047	0.041	0.104	1.145	0.253
Experience	–0.087	0.035	–0.216	–2.456	0.015*

Source: Primary data.

Note: * significant at 0.05.

Table 5. Model Summary.

Model	*R*	*R*-squared	Adjusted *R*2	ANOVA (F)	Sig.
1	0.652	0.424	0.418	68.84	0.000*
2	0.700	0.489	0.475	6.99	0.000*

Source: Primary data.

Note: * significant at 0.05.

Table 6. Demographic Variable-based Variations in Occupational Stress.

Demographic Variables	Categories	Mean	*t*-value/*F*- value	Sig.
Gender	Male	3.96	2.207	0.029
	Female	3.80		
Marital status	Married	3.89	–1.339	0.182
	Unmarried	4.00		
Age	21–31 years	3.99	12.545	0.000*
	32–42 years	4.09		
	43–53 years	3.56		
	54–64 years	3.40		
Educational qualification	Postgraduate	4.07	7.319	0.000*
	Graduate	4.00		
	Intermediate	3.83		
	Matriculation	3.47		

(Table 6 continued)

(Table 6 continued)

Demographic Variables	Categories	Mean	*t*-value/*F*- value	Sig.
Job experience	2–10 years	4.03	15.072	0.000*
	11–19 years	4.06		
	20–28 years	3.62		
	29–37 years	3.33		
Rank	Inspector	4.03	10.037	0.000*
	Sub-inspector	3.94		
	Assistant sub-inspector	3.79		
	Head constable	3.30		
	Constable	3.05		
Grade	Grade I	4.00	–4.923	0.000*
	Grade II	3.34		

Source: Primary data.

Note: * significant at 0.05.

Table 7. Demographic-variable Based Regression

Variable	Category	Regression Equation	*R*-square
Age	21–31	OS = 5.177+ (–0.398)x	0.334
	32–42	OS = 4.636+(–0.212)x	0.060
	43–53	OS = 5.292+(–0.529)x	0.382
	54–64	OS = 5.417+(–0.604)x	0.478
Educational qualification			
	Postgraduate	OS = 5.514+(–0.513)x	0.369
	Graduate	OS = 5.112+(–0.376)x	0.323
	Intermediate	OS = 4.566+(–0.518)x	0.318
	Matriculation	OS = 5.195+(–0.266)x	0.071
Job experience			
	2–10	OS = 5.208+(–0.395)x	0.321
	11–19	OS = 5.203+(–0.431)x	0.324
	20–28	OS = 3.685+(–0.056)x	0.003
	29–37	OS = 5.481+(–0.611)x	0.441
Rank	Inspector	OS = 4.278+(–0.235)x	0.294
	Sub-inspector	OS = 4.331+(–0.616)x	0.568
	Assistant sub-inspector	OS = 5.336+(–0.474)x	0.351
	Head constable	OS = 5.163+(–0.397)x	0.235
	Constable	OS = 4.948+(–0.328)x	0.172
Grade	Grade I	OS = 5.035+(–0.517)x	0.492
	Grade II	OS = 4.992+(–0.343)x	0.188

Source: Primary data.

Discussion

The present study has three major findings. First, WPS has a significant negative association with OS. Second, stress increases with the rank, grade and educational qualifications of the employees. Also, stress initially increases with age and experience of the police personnel; after that, it decreases. Third, WPS is a more significant predictor of OS for employees who are more stressed. These results could be explained in the light of the job requirement of policing and Indian culture.

One of the dimensions of WPS, meaningful work, reports a significant positive correlation with OS. The regression coefficient was also positive but insignificant. It means meaningful job assignments might lead to stressed police personnel. The finding is not in line with previous revelations but could be due to the nature of the job of police personnel. According to Smith & Charles (2010), employees who engaged in highly meaningful activities and contributed to greater causes could not escape being the victims of job stress. The results could be understood in the light of the job demands–resource (JDR) model. Although meaningful work is job positive (resource), it is accompanied by higher job demands for police personnel. A significant job for a police officer includes investigating a crime, forecasting and preventing crime and untoward incidents, developing a strategy to capture a culprit and coordinating with different agencies like the National Investigation Agency (NIA), Central Bureau of Investigation (CBI), police department of other states, etc. Routine jobs includes issuing traffic challans, guarding VIP movements, paper work, etc. It is very clear from the nature of work that meaningful work demands higher involvement and is psychologically draining. A higher amount of risk and uncertainty is also associated with such significant assignments. Thus, meaningful work may lead to an imbalance of resources and demands for police officers, leading to higher stress levels.

The findings also indicate that sense of community could explain significant variance in occupational stress. As individuals spend maximum time at the workplace, co-workers become relatively more important than their relatives and social groups. Hence positive relations at the workplace foster a positive sense of community. However, rewards and promotions may result in unhealthy competition, internal rivalry and a lack of compassion for fellow workers. It may create an environment of distrust, high individualism and low connection amongst the employees. Facilitating connectedness amongst the employees on mental, emotional and spiritual levels (Neal & Bennett, 2000) will improve the compassion and nurture harmony, which will help avoid distress.

Moreover, embracing spirituality does not discourage competition; it discourages unhealthy competition (Dandona, 2013). The improved attachment would encourage the individuals to discuss their work problems and lighten their burdens with their colleagues, thereby reducing levels of occupational stress. Sharma (2009) states that sometimes, all they desire is a shoulder to discuss their woes in a stressful occupation like policing. An organisational support system is imperative for a motivated workforce which is relatively weak in Indian police organisations. Lack of support from superiors results in low morale and a stressed

workforce (Tyagi & Dhar, 2014). Khadse et al. (2020) suggested developing 'peer support networks' in the police department. Henceforth, promoting community will promote spirituality on a group level in organisations and aid in controlling work stress.

Next, it is reported that alignment with organisational values also explains significant variance in occupational stress. Incongruity in the value system contributes to increased anxiety amongst the employees (Mohanty, 2017). In police organisations, employees often face a disagreement between the organisation's and their values (e.g., setting a criminal free on the grounds of insufficient evidence). Proper orientation and adequate training in the organisations can help build a strong alignment between the value system of the employees and organisations (Umranikar, 2009). Moreover, senior and experienced employees can share their experiences about conflicting situations and their approaches to avoid distress and mental unrest. Enhanced agreement of values will help avoid distress and guide the employees' actions during crucial decisions or emergencies.

The study also found that WPS is a relatively stronger predictor of OS for more stressed employees. In simple words, WPS works more effectively for highly stressed police personnel. Similar findings were reported by Saxena et al. (2020). They reported that WPS has a more substantial effect on OS for offshore employees of Indian oil and gas companies.

Theoretical Implications of the Study

The present study has enriched existing literature on WPS and OS in three possible ways. First, meaningful work, a sense of community and alignment with organisational values are the three most widely used dimensions of WPS. Several previous researchers have relied on these three constructs to explore the utilities of WPS for modern workplaces empirically. The present paper elaborates on these WPS dimensions from an Indian perspective. In other words, this paper is an attempt towards the Indian conceptualisation of WPS. Second, the study offers a counterintuitive finding, that is, a positive correlation between meaningful work and OS. A theoretical understanding and previous studies have reported a negative linkage between the two variables; however, the present study offers something different. The finding highlights the need to examine the nature of the job before proposing any spiritual intervention for any organisation. It indicates that the adoption of a tailormade WPS policy for a specific job profile. Thus, future researchers may find it interesting to investigate the impact of WPS on OS in other vital occupations such as medical professionals, military personnel, firefighters, journalists, etc. Third, the study concludes that WPS works better for more stressed employees. Although the generalisation requires more confirmatory studies, the finding is unique and novel. It may encourage future researchers to investigate similar phenomena in other settings too.

Practical Implications of the study

The findings present significant implications for organisations striving to reduce the level of stress faced by their employees. Specifically, organisations need to imbibe the spiritual aspect in various processes of their functioning. For instance, the current training modules in the organisations primarily focus on developing their employees' physical and mental abilities with negligible attention on their spiritual outlook (Moran, 2017). Yoga sessions have been increasingly adopted as a stress-busting mechanism to promote physical and mental wellness. There is little attention on its contribution to the spiritual development of an individual. A holistic attitude to comprehend the immense potential benefits of various spiritual practices needs to be considered to develop an effective people-oriented organisation. Additionally, human resource development programs may consider developing the spirituality quotient (SQ) of their employees. It can also be used for selection, training needs evaluation, work team formation, etc. Moreover, effective teams can act as a valuable source of connection and social support to the employees, which will help lighten the emotional labour faced in stressful occupations.

The findings of the present study may help in designing any spiritual intervention for police officers. As stress increases with the designation and grade of employees, senior police officers may be given more intense spiritual training programs. Police academy might develop their own spiritual trainers or take help from other spiritual organisations like The Art of Living, Isha foundation, etc. Trainees of Indian Administrative Officers are offered the 'Inner Engineering Program'. This program helps future bureaucrats to learn the perfect alignment of body, mind and spirit. A similar sort of training may be arranged for senior police officers. The junior officers may be required to attend various seminars and workshops on spiritual practices like meditation, *pranayam*, self-introspection, etc. Police personnel may also be provided with books and lectures on ancient Indian notions like *karma yoga, swadharma*, and *bhakti-yoga*.

The learnings from local Himachali culture could also help in designing *Pahari*-specific spiritual interventions for police officers. Himachal Pradesh is also referred to as 'Dev Bhoomi', which means 'the abode of the Gods'. The *Pahari* communities have their local deities called 'Devtas' (Mehta, 2020). It is believed that these deities reside in water bodies, fire, sky, earth, valleys and mountains, stones and rocks and forests (Sharma, 2009). Such belief brings a lot of respect and devotion to nature and natural artifacts. In this background, spiritual interventions like nature walking, *yagna* or *havan* (fire ritual), water festivals such as *Holi*, and prayers designated to mountains and rocks like *Govardhan puja* (festival worshiping Govardhan hill) could be highly beneficial. These interventions are derived from their local culture; thus, they may significantly impact local police officers. Also, local *Pahari* people are well known for their life contentment and gratefulness. Police academies may apply different gratitude interventions such as 'gratitude letters', 'count your blessings', 'self-introspection', etc., to strengthen their gratitude traits.

Limitations of the Study and Conclusion

Apart from presenting significant implications, the present study also has certain limitations. First, spirituality is believed to emerge from within an individual, making it a highly personal construct. Using a survey instrument to measure the dimension poses several challenges (Milliman et al., 2003). Second, the responses have been collected based on self-report data, which suffers from its limitations and biases. Third, the present study does not consider the influence of potential moderating variables such as spiritual leadership, personality traits, etc. Future studies involving moderating variables can offer unique perspectives on the relationship between the study variables. The present study provides empirical evidence on the significant association between workplace spirituality and occupational stress, wherein incorporating spirituality at the workplace would decrease stress levels. The findings have important implications for the police organisation and other workplaces that are continuously battling the menace of work stress in the present era of mindless globalisation and deteriorating ethics. To advance as a spiritually oriented workplace, the spiritual element needs to be imbibed in the organisation's work culture and core values. Moreover, nurturing spiritual energy is an effective strategy and an obligation for the employee well-being and service-oriented organisations.

Annexure: Factor Structure of WPS Scale.

Items	Loadings	*H*2	Eigen Value
Meaningful work			9.208
03 Experience joy in my work.	0.647	0.607	
04 Spirit is energised by my work.	0.669	0.593	
07 Work is connected to what I think is important in life	0.751	0.692	
10 Look forward to coming to work.	0.600	0.535	
13 See a connection between work and social good	0.654	0.607	
15 Understand what gives my work personal meaning	0.713	0.697	
Sense of Community			1.942
01 Believe people support each other	0.708	0.757	
02 Working cooperatively with others is valued	0.693	0.771	
05 Feel free to express opinions.	0.564	0.379	
06 Feel part of a community	0.725	0.579	

(Annexure continued)

(Annexure continued)

Items	Loadings	*H*2	Eigen Value
08 Think employees are linked with a common purpose.	0.766	0.613	
19 Believe employees genuinely care about each other	0.630	0.524	
20 Feel there is sense of being part of family	0.492	0.502	
Alignment with organisational values			1.760
09 Feel connected with the organisational goals	0.789	0.691	
11 Organisation cares about all its employees.	0.657	0.645	
12 Organisation is concerned about the health of employees.	0.616	0.603	
14 Feel connected with the mission of my organisation.	0.774	0.661	
16 Feel positive about the values of the organisation.	0.762	0.637	
17 Organisation is concerned about the poor.	0.702	0.549	
18 Organisation cares about whether my spirit is energised by my work.	0.628	0.693	
21 Organisation has a conscience	0.668	0.573	
Number of items		21	
Eigen values	12.91		

Source: Primary data.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The authors received no financial support for the research, authorship and/or publication of this article.

ORCID iDs

Stuti Jalan https://orcid.org/0000-0002-9059-6543

Naval Garg https://orcid.org/0000-0001-6694-9663

References

Almale, B. D., Vankudre, A. J., Bansode-Gokhe, S. S., & Pawar, V. K. (2014). An epidemiologic study of occupational stress factors in Mumbai police personnel. *Indian Journal of Occupational and Environmental Medicine*, *18*(3), 109–112. https://doi.org/10.4103/0019-5278.146906

Anderson, G., Litzenberger, R., & Plecas, D. (2002). Physical evidence of police officer stress. *Policing: An International Journal of Police Strategies and Management*, *25*(2), 399–420. https://doi.org/10.1108/13639510210429437

Arnetz, B. B., Ventimiglia, M., Beech, P., DeMarinisg, V., Lökkh, J., & Arnetz, J. E. (2013). Spiritual values and practices in the workplace and employee stress and mental well-being. *Journal of Management, Spirituality, & Religion*, *10*(3), 271–281. https://doi.org/10.1080/14766086.2013.801027

Ashmos, D. P., & Duchon, D. (2000). Spirituality at work: A conceptualization and measure. *Journal of Management Inquiry*, *9*(20), 134–145. https://doi.org/10.1177/105649260092008

Bell, A., Rajendran, D., & Theiler, S. (2012). Spirituality at work: An employee stress intervention for academics. *International Journal of Business and Social Science*, *3*(11), 68–82.

Bhawuk, D. P. S. (2011). *Spirituality and Indian psychology—Lessons from the Bhagvad Gita*. Springer Science+Business Media.

Bosch, L. (2009). The inevitable role of spirituality in the workplace. *Business Intelligence Journal*, *2*(1), 139–157.

Cartwright, S., & Holmes, N. (2006). The meaning of work: The challenge of regaining employee engagement and reducing cynicism. *Human Resource Management Review*, *16*(2), 199–208.

Chand, P., & Koul, H. (2012, March 17–18). *Workplace spirituality, organizational emotional ownership, and job satisfaction as moderators in coping with job stress*. International Conference on Humanities, Economics, and Geography. http://psrcentre.org/images/extraimages/2.%20312133.pdf

Colligan, T. W., & Higgins, E. M. (2005). Workplace stress: Etiology and consequences. *Journal of Workplace Health*, *21*(2), 90–97.

Cooper, C. L., & Marshall, J. (1976). Occupational sources of stress: A review of the literature relating to coronary heart diseases and mental health. *Journal of Occupational Psychology*, *49*(1), 11–28. https://doi.org/10.1111/j.2044-8325.1976.tb00325.x

Cooper, C. L., Dewe, P. J., & O'Driscoll, M. P. (2001). *Organizational stress: A review and critique of theory, research, and applications*. SAGE Publications.

Cox, C. J. (2011). *The relationship between stress, spirituality, and academic performance* [Doctoral dissertation, Bowling Green State University]. OhioLINK Electronic Thesis and Dissertation Centre.

Csiernik, R., & Adams, D. (2012). Spirituality, stress, and work. *Employee Assistance Quarterly*, *18*(2), 29–37.

Cunningham, C. J. L. (2014). *The role of demographics in occupational stress and well being*. Emerald Group Publishing Limited.

Dandona, A. (2013). Spirituality at workplace and job satisfaction. *International Journal of BioSciences, Alternative and Holistic Medicine*, *4*(1), 1–9.

Daniel, J. L. (2015). Workplace spirituality and stress: Evidence from Mexico and US. *Management Research Review*, *38*(1), 29–43.

Dehler, E. G., & Welsh, A. M. (2003). The experience of Work- Spirituality and the new workplace. In R. A. Giacalone & C.L. Jurkiewicz (Eds), *The handbook of workplace spirituality and organizational performance*. M. E. Sharper.

Dehler, G., & Welsh, M. (1994). Spirituality and organizational transformation: Implications for the new management paradigm. *Journal of Managerial Psychology*, *9*(6), 17–26.

Dehler, G., & Welsh, M. (2010). The experience of work: Spirituality and the new workplace. In R. A. Giacalone & C. L. Jurkiewicz (Eds), *Handbook of workplace spirituality and organizational performance* (pp. 59–72). M.E. Sharper, Inc.

Edwards, J. R. (1992). A cybernetic theory of stress, coping, and well-being in organizations. *Academy of Management Review*, *17*(2), 238–274.

Eisler, R., & Montouri, A. (2002). The human side of spirituality. In Giacalone, R. A., & Jurkiewicz, C. L. (Eds), *Handbook of workplace spirituality and organizational performance* (pp. 46–56). M.E. Sharper, Inc.

Folkman, S. (1997). Positive psychological states and coping with severe stress. *Social Science & Medicine*, *45*(8), 1207–1221. https://doi.org/10.1016/s0277-9536(97)00040-3

Follet, M. P. (1930). *Creative experiences*. Green & Company.

French, J. R. P., Caplan, R. D., & Harrison, R. V. (1982). *The mechanisms of job stress and strain*. Wiley Books.

Garcia-Zamor, J. (2003). Workplace spirituality and organizational performance. *Public Administration Review*, *63*(3), 355–363.

Garg, N. (2017). Workplace Spirituality and Employee Well-being: An Empirical Exploration. *Journal of Human Values, 23*(2). https://doi.org/10.1177/097168581668974

Garg, N., Punia, B. K., & Jain, A. (2019). Workplace spirituality and job satisfaction: Exploring mediating effect of organization citizenship behaviour. *Vision: The Journal of Business Perspective*, *23*(3), 287–296.

Giacalone, R. A., & Jurkiewicz, C. L. (2003). Toward a science of workplace spirituality. In R. A. Giacalone & C. L. Jurkiewicz (Eds), *The handbook of workplace spirituality and organizational performance*. M E Sharper.

Goyal, M., Singh, S., Sibinga, E. S., Gould, N. F., Seymour, A. R., Sharma, R., Berger, Z., Sleicher, D., Maron, D., Shihab, H. M., Ranasinghe, P. D., Linn, S., Saha, S., Bass, E. B., & Haythornthwaite, J. A. (2014). Meditation programs for psychological stress and well-being: A systematic review and meta-analysis. *JAMA Internal Medicine*, *174*(3), 357–368.

Holmlund-Rytkönen, M., & Strandvik, T. (2005). Stress in business relationships. *Journal of Business & Industrial Marketing*, *20*(1), 12–22.

Jnaneshwar, K., & Sulphey, M. M. (2021). A study on the relationship between workplace spirituality, mental well-being and mindfulness. *Management Science Letters*, *11*(3), 1045–1054.

Kahn, R. L., Wolfe, D. M., Quinn, R. P., Snoek, J. D., & Rosenthal, R. A. (1964). *Organizational stress: Studies in role conflict and ambiguity*. Wiley.

Karakas, F. (2010). Spirituality and performance in organizations: A literature review. *Journal of Business Ethics*, *94*(1), 89–106.

Karasek, R. A., Triantis, K. P., & Chaudry, S. S. (1982). Co-worker and supervisor support as moderators of associations between task characteristics and mental strain. *Journal of Occupational and Behavior*, *3*(2), 181–200.

Kelloway, E. K., Teed, M., & Kelly, E. (2008). The psychosocial environment: Towards an agenda for research. *International Journal of Workplace Health Management*, *1*(1), 50–64.

Kenny, D. T., & Cooper, C. (2003). Introduction: Occupational stress and its management. *International Journal of Stress Management*, *10*(4), 275–279.

Khadse, P. A., Gowda, G. S., Ganjekar, S., Desai, G., & Murthy, P. (2020). Mental health impact of COVID-19 on police personnel in India. *Indian Journal of Psychological Medicine*, *42*(6), 580–582.

Kline, R. B. (2005). *Principles and Practice of Structural Equation Modelling* (2nd ed.), New York: Guildford Press.

Kolodinsky, R., Giacalone, R., & Jurkiewicz, C. (2008). Workplace values and outcomes: Exploring personal, organizational, and interactive workplace spirituality. *Journal of Business Ethics*, *81*(2), 465–480.

Kumar, V., & Kumar, S. (2014). Workplace spirituality as a moderator in relation between stress and health: An exploratory empirical assessment. *International Review of Psychiatry*, *26*(3), 344–351.

Lal, R., Pathak, P., & Chaturvedi, K. R. (2017). A study of spiritualistic techniques of stress among engineering student: With special reference of assessment of perceived stress and spirituality. *Indian Journal of Positive Psychology*, *8*(3), 333–337.

Landis, B. J. (2009). Uncertainty, spiritual well-being, and psychosocial adjustment to chronic illness. *Issues in Mental Health Nursing*, *17*(3), 217–231.

Lazarus, R. S. (1991). Psychological stress in the workplace. *Journal of Social Behavior and Personality*, *6*(7), 1–13.

Maheshwari, N. (2015). Spiritual intelligence: Occupational commitment. *SCMS Journal of Indian Management*, *12*(2), 29–38.

Mahipalan, M., & Suresh, S. (2019). Workplace spirituality, psychological well-being, and mediating role of subjective stress: A case of secondary school teachers in India. *International Journal of Ethics and Systems*, *35*(4), 725–739.

Martinussen, M., & Richardsen, A. M. (2006). Air traffic controller burnout: Survey responses regarding job demands, job resources, and health. *Aviation, Space & Environmental Medicine*, *77*(4), 422–428.

Maslach, C., & Leiter, M. P. (1997). *The truth about burnout*. Jossey-Bass.

Mathur, P. (1995). Perceptions of police stress: An empirical study of stressors and coping responses among police personnel in India. *Indian Journal of Criminology*, *23*(1), 9–19.

McMillan, D. W., & Chavis, D. M. (1986). Sense of community: A definition and theory. *Journal of Community Psychology, 14*(1), 6-23.

Mehta, A. (2020). Eco-critical reading of Devta Culture: Himachal Pradesh. Melange Conference. Retrieved from https://www.researchgate.net/publication/339681314_Eco-critical_reading_of_Devta_Culture_Himachal_Pradesh

Michie, S. (2002). Causes and management of stress at work. *Journal of Occupational and Environmental Medicine*, *59*(1), 67–72.

Milliman, J., Czaplewski, A. J., & Ferguson, J. (2003). Workplace spirituality and employee work attitudes. An exploratory empirical assessment. *Journal of Organizational Change Management*, *16*(4), 426–447.

Mishra, A. K., & Bisht, S. (2014). Spiritual orientation and perceived stress level: An empirical study on Indian youth. *IPE Journal of Management*, *4*(2), 119–130.

Mitroff, I., & Denton, E. (1999). A study of spirituality in the workplace. *Sloan Management Review*, *40*(4), 83–91.

Mohanty, S. P. (2017). *Impact of police culture in the lives of police personnel: A study on the police commissionerate of Bhubaneswar & Cuttack* [Doctoral dissertation, KIIT University]. https://shodhganga.inflibnet.ac.in/handle/10603/190613

Moran, R. (2017). Workplace spirituality in law enforcement: A content analysis of the literature. *Journal of Management, Spirituality, & Religion*, *14*(4), 343–364. https://doi.org/10.1080/14766086.2017.1376287

Naylor, T., Willimon, W., & Osterberg, R. (1996). The search for community in the workplace. *Business and Society Review*, *97*, 42–47.

Neal, J. (2013). Faith and spirituality in the workplace: Emerging research and practice. In J. Neal (Ed.), *Handbook of faith and spirituality in the workplace: Emerging research and practice* (pp. 3–5). Springer Science+BusinessMedia. https://doi.org/10.1007/978-1-4614-5233-1_1

Neal, J. A., & Bennett, J. (2000). Examining multi-level or holistic spiritual phenomena in the workplace. In D. C. Trott (Ed.), *Management, spirituality, & religion newsletter* (pp. 1–2). Academy of Management.

Nikam, K. P., & Shaikh, M. (2014). Occupational stress, burnout, and coping in police personnel: Findings from a systematic review. *American International Journal of Research in Humanities, Arts and Social Sciences*, *6*(2), 144–148.

Noor, S., & Arif, S. (2011). Achieving job satisfaction via workplace spirituality: Pakistani doctors in focus. *European Journal of Social Sciences*, *19*(4), 507–515.

Palmer, S., Cooper, C., & Thomas, K. (2004). A model of work stress to underpin the health & safety executive advice for tackling work-related stress and stress risk assessments (pp. 2–5). *Counselling at Work, Winter*.

Pande, N., & Naidu, R. K. (1992). Anasakti and health: A study of non-attachment. *Psychology and Developing Societies*, *4*(1), 89–104.

Pandey, A., Gupta, R. K., & Kumar, P. (2016). Spiritual climate and its impact on learning in teams in business organizations. *Global Business Review, 17*(3). https://doi.org/10.1177/0972150916631208

Pandey, A., Gupta, R., & Arora, P. A. (2009). Spiritual climate of business organizations and its impact on customers' experience. *Journal of Business Ethics*, *88*, 313–332.

Pandey, A., Gupta, V., & Gupta, R. K. (2019). Spirituality and innovative behaviour in teams: Examining the mediating role of team learning. *IIMB Management Review*, *31*(2), 116–126.

Pandya, S. P. (2017). Dealing with crime, criminals and the criminal justice system through spirituality. *Police Journal, 90*(4), 348–377.

Pardasani, R., Sharma, R., & Bindlish, P. (2014). Facilitating workplace spirituality: Lessons from Indian spiritual traditions. *Journal of Management Development*, *33*(8/9), 847–859.

Pareek, U. (1983). *Organizational role stress scale manual*. Navina Publications.

Paton, D., & Violanti, J. M. (1996). *Traumatic stress in critical occupations—Recognition, consequences & treatment*. Charles C. Thomas Publication.

Pawar, B. S. (2009). Individual Spirituality, workplace spirituality, and work attitudes—An empirical test of direct and interaction effects. *Leadership & Organization Development Journal*, *30*(8), 759–777.

Pestonjee, D. M., & Mishra, P. K. (1999). Role stress and job satisfaction amongst doctors. *Journal of Health Management*, *1*(1), 117–131.

Pfeffer, J. (2003). Business and spirit: Management practices that sustains values. In R. A. Giacalone & C. L. Jurkiewicz (Eds), *Handbook of workplace spirituality and organizational performance* (pp. 29–45). M.E. Sharper, Inc.

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology, 63*, 539–569. https://doi.org/10.1146/annurev-psych-120710-100452

Poe, R., & Courter, C. L. (1995). Spirituality at work. *Across the Board*, *32*(5).

Powell, L. H., Shahabi, L., & Thoresen, C. E. (2003). Religion and spirituality: Linkages to physical health. Religion and Spirituality. *American Psychologist*, *58*(1), 36–52. https://doi.org/10.1037/0003-066X.58.1.36

Rathee, R., & Rajain, P. (2020). Workplace spirituality: A comparative study of various models. *Jindal Journal of Business Research*, *9*(1), 27–40.

Ratna, R., Chawla, S., & Mittal, R. (2013). Organizational role stress: Level of stress, major stressor, and its differences. *International Journal of Indian Culture and Business Management*, *7*(3), 359–383.

Rego, A., & Cunha, M. P. E. (2008). Workplace spirituality and organizational commitment: An empirical study. *Journal of Organizational Change Management*, *21*(1), 53–75.

Robertson, I., & Cooper, C. (2011). *Well-being: Productivity and happiness at work*. Palgrave Macmillan.

Saini, R. (2017). Developing spirituality at workplace: Boon or bane. *Indian Journal of Occupational & Environmental Medicine*, *21*(2), 43–44.

Sarkar, A. & Garg, N. (2020). "Peaceful workplace" only a myth? Examining the mediating role of psychological capital on spirituality and nonviolence behaviour at the workplace. *International Journal of Conflict Management, 31*(5). https://doi.org/10.1108/IJCMA-11-2019-0217

Savery, L., Soutar, G., & Weaver, J. (1993). Stress and the police officer: Some West Australian evidence. *The Police Journal: Theory, Practice, and Principles*, *6*(3), 277–290.

Saxena, A., Garg, N., Punia, B. K., & Prasad, A. (2020). Exploring role of Indian workplace spirituality in stress management: A study of oil and gas industry. *Journal of Organizational Change Management*, *33*(5), 779–803.

Schneiders, S. M. (2008). Religion vs. spirituality: A contemporary conundrum. *Spiritus*, *3*(2), 163–185. Reprinted in *St. Augustine Papers*, *9*(1), 23–53.

Sharafi, M. R., Hoshmandja, M., & Poorkhalil, M. (2014). The role of organizational spirituality in job stress and job satisfaction in staff of Baqiyatallah Hospital personnel. *Journal of Educational and Management Studies*, *4*(4), 885–890.

Sharma, K. K. (1985). *Law & order administration*. National Book Organisation.

Sharma, P. (2009). Occupational stress & mental health. *The Indian Police Journal*, *66*(3), 15–18.

Smith, J., & Charles, G. (2010). The relevance of spirituality in policing: A dual analysis. *Journal of Police Science & Management*, *12*(3), 320–338.

Spielberger, C. D., Westberry, L. G., Grier, K. S., & Greenfield, G. (1981). *The police stress survey: Sources of stress in law enforcement*. Monograph. Human Resources Institute.

Srirangarajan, G. S., & Bhaskar, R. K. (2011). Key dimensions of spirit at work: An Indian perspective. *Journal of Human Values*, *17*(2), 93–120.

Suresh, V., & Janaki, T. (2012). Dharma, quality and social responsibility: An imperative linkage to effective corporate governance. *EXCEL International Journal of Multidisciplinary Management Studies*, *2*(3), 78–90.

Sutherland, V. J., & Cooper, C. L. (2000). *Strategic stress management*. Palgrave Macmillan. https://doi.org/10.1057/9780230509146_7

Tischler, L., Biberman, J., & McKeage, R. (2002). Link emotional intelligence, spirituality, and workplace performance: Definitions, model, and ideas for research. *Journal of Managerial Psychology*, *17*(3), 203–218.

Tyagi, A., & Dhar, L. R. (2014). Factors affecting health of the police officials: Mediating role of job stress. *Policing: An International Journal of Police Strategies & Management*, *37*(3), 649–664.

Umranikar, J. Y. (2009). *Police reforms in India, a Sisyphean saga: Functional review of India Police*. Ameya Prakashan.

Usman, A., & Danish, R. Q. (2010).Spiritual consciousness in banking managers and its impact on job satisfaction. *International Business Research*, *3*(2), 65–72.

Wu, A. M. S., & Wong, E. M. W. (2008). Disordered gambling among Chinese casino employees. *Journal of Gambling Studies*, *24*(2), 207–217.

Zellars, K., & Perrewe, K. (2003). The role of spirituality in occupational stress and well-being. In R. A. Giacalone & C. L. Jurkiewicz (Eds), *Handbook of workplace spirituality and organizational performance* (pp. 300–305). M.E. Sharper, Inc.

Zou, W. C., & Dahling, J. (2017). Workplace spirituality buffers the effects of emotional labour on employee well-being. *European Journal of Work and Organizational Psychology*, *26*(5), 768–777.

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)
Pune, India. Oct 29-30, 2021

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON) | 978-1-6654-2503-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/SMARTGENCON51891.2021.9645798

FPGA Based Comparative Analysis Of Single Carrier PWM Schemes

Yatharth Gupta
Department Of Electrical Engineering
Delhi Technological University
New Delhi, India
yatharthst6@gmail.com

Madhusudan Singh
Department Of Electrical Engineering
Delhi Technological University
New Delhi, India
madhusudan@dce.ac.in

***Abstract*—The purpose of this work is to design, simulate and generate Verilog codes for digital PWM techniques for sinusoidal pulse width modulation and third-harmonic PWM injection with different carriers for a three-level voltage source inverter using Xilinx system generator. The built-in FPGA architecture to generate PWM pulses is used.**

Keywords—pulse width modulation control(PWM),Xilinx System Generator(XSG), Inverter, field programmable gate array(FPGA).

I. INTRODUCTION

In the past few years, there has been an exponential increase in pace of devlopment of digital controllers. Low frequency applications need for pulse width modulation is catered by micrcontroller[1] and digital signal processors(upto 10Khz). Due to the advantages of parallel processing and high frequency of operation FPGA's (upto 3MHz) are preferred for realizing less power applications.

Pulse width modulation (PWM) is a popular technique in which the duty cycle of a pulse signal is used to depict the magnitude of a reference signal. The intersection of the reference voltage curve and the carrier waveform[2-3] determines the duty cycle of the switch.It is essentially a square pulse with variable width, span of which can be adjusted as required. Primarily used for voltage control, PWM techniques are efficient and widely used to control the drives of switching devices. It is used to reduce power loss especially when it is made of purely resistive loads..For motor speed control, the modulated duty cycle is used to control speed. Compared with the turn-off time, the longer the turn-on time, the more power is provided to the load.

A single PWM method may not be suitable for all applications[3].

In this paper the PWM schemes compared are:
1)Sine PWM
2)Third Harmonic Injection PWM.
Two different types of carrier waves are:
1)Sawtooth wave.
2)Triangluar wave.

II. SYSTEM DESIGN

Figure 1 shows a general system consisting of a three-stage three-phase inverter with 6 switches and a three-phase load (consisting of L and R elements). The system description is as follows:

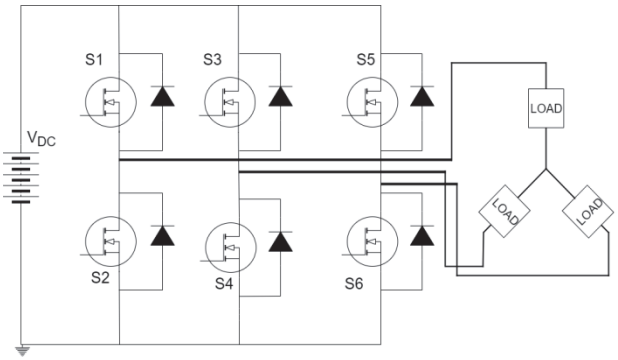


Fig.1. Three phase Voltage Source Inverter

TABLE I. SYSTEM DESCRIPTION

System Specifications	
DC Bus Voltage	300V
Reference Wave Fundamental Frequency	50Hz
Carrier Wave frequency	10KHz
Modulation Index	0.999
Load Resistance	200Ω
Load Reactance	24mH

Same system parameters are used for both the PWM methods

A. Sine Pulse Width Modulation

This method involves, the Sine reference wave being compared with the high-frequency carrier wave.A comparator is used to compare their respective magnitudes. The output pulse is digital high whenever the magnitude of the reference wave is greater than the carrier wave, else it is zero. A high output implies that the switch is open, and a low output means the switch is closed. For a three-phase system, the reference wave is offset by 120° and 240° from the original single-phase wave and compared with the carrier to generate pulses.

Generally, this method results in an output voltage with less harmonic distortion. However, the six-step mode for unit magnitude of the reference wave outputs the magnitude of the fundamental wave for unit magnitude of refer which is 78.54% of the original wave.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:59:32 UTC from IEEE Xplore. Restrictions apply.

When the modulation index m is increased beyond unity, to extend the output beyond 78.54% of the maximum value,the effect of the harmonics of 3rd,9th,15th,21st order result in vanishing of pulses.This is due to the elimination of these harmonics through the three-phase topology.

B. *Third Harmonic Injection PWM*

The operating principle of THIPWM consists of the addition of a fundamental component with addition of third harmonic component:.

i.e

$$y(t) = K_1 sin(\omega t) + K_2 sin(3\omega t) \quad (1)$$

Where K_1 and K_2 are optimized coefficients to get the requiste output.Comparing the differentiation of the above equation to zero gives the optimized values of the constants as,

K_1=1.156 and K_2=0.1924 respectively.

Therefore,the equations for 3-phases are given by:

$$V_a(t) = 1.156 * sin(\omega t) + 0.1924 * sin(3\omega t) \quad (2)$$

$$V_b(t) = 1.156 * sin(\omega t - 2\Pi 3) + 0.1924 * sin(3\omega t) \quad (3)$$

$$V_c(t) = 1.156 * sin(\omega t + 2\Pi 3) + 0.1924 * sin(3\omega t) \quad (4)$$

Identical to SPWM,in THIPWM also pulses are generated by comparison of carrier wave with the new reference wave.For the same supply voltage and magnitude of fundamental wave the reference wave in THIPWM is 15.5% bigger than the reference wave of SPWM,resulting in roughly 90% of the amplitude than obtained by a six-step mode.

III. FPGA DESIGN

A. *Generation Of Reference Wave*

A ROM block is used to generate a 50 Hz sine wave and is connected to a fixed counter as shown(below in figure 2). The Xilinx ROM module is a read-only memory (ROM) with a single port. Appropriate parameters such as depth, word type, number of bits and number of decimal places is adjusted to obtain a fundamental wave output.

To avoid dead band due to delay induced 3phase wave generation, the trigonometric expansion of the phase displaced waves are used to obtain the required waves.

i.e

$$sin(A + B) = sin(A)cos(B) - cos(A)sin(B) \quad (5)$$

In order to obtain the THIPWM signal, the required gain blocks and adders are used to obtain the required carrier wave.

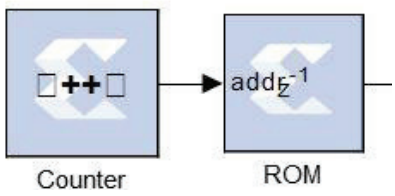


Fig.2. Generation Of Sine Wave

B. *Generation Of Carrier Waves*

Different types of up,down or up/down counter which can be inexhaustible or of limited count can be implemented by the Xilinx counter module(as shown in figure 3). The datatype of unsigned fixed point number is used.

To obtain different types of carrier waves :

1) Leading edge carrier wave:

A limited range up counter is used to generate a leading edge carrier wave.

Fig. 3: Generation Of Saw-tooth Carrier Wave

2) Triangular Carrier Wave:

Xilinx blocksets such as absolute, multiplier, adder are used to obtain the triangular modulated carrier signal(as shown below in figure 4).

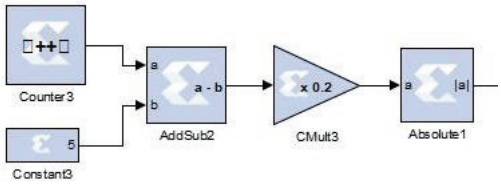


Fig. 4: Generation Of Triangular Carrier Wave

C. *Generation Of Pulses*

The adder and comparator blocks are used to compare the different carrier wave with the corresponding reference wave to generate the pulse width modulation pulses[6].

The generated pulses are fed to the three-phase voltage source inverter and the results are recorded.

IV. RESULTS AND DISCUSSIONS

A. *Sine Carrier Wave:*

The obtained carrier wave for three phase sine wave (as shown below in figure 5):

Fig.5. Three Phase Sinusoidal Reference Wave

B. *Reference wave for three phase THIPWM (is shown below in figure 6).*

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:59:32 UTC from IEEE Xplore. Restrictions apply.

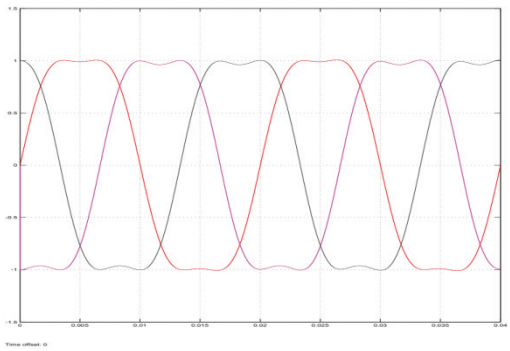


Fig. 6. Reference Wave for Third Harmonic Injected Carrier

C. The different carrier waves are shown below:

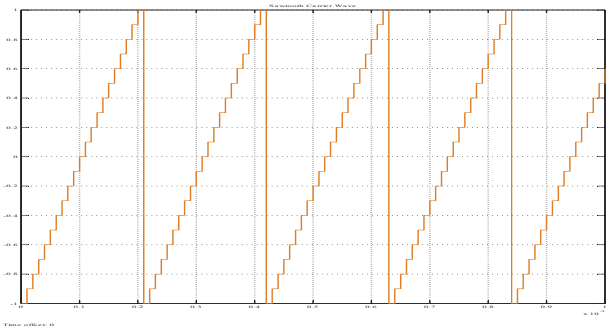


Fig.7. Sawtooth Carrier Wave

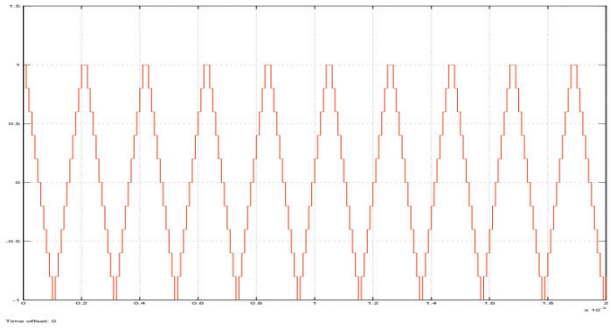


Fig. 8. Triangular Carrier Wave

Figure 7 and figure 8 above show saw-tooth and triangular carrier wave respectively.

D. The THD obtained in the various configurations are shown below:

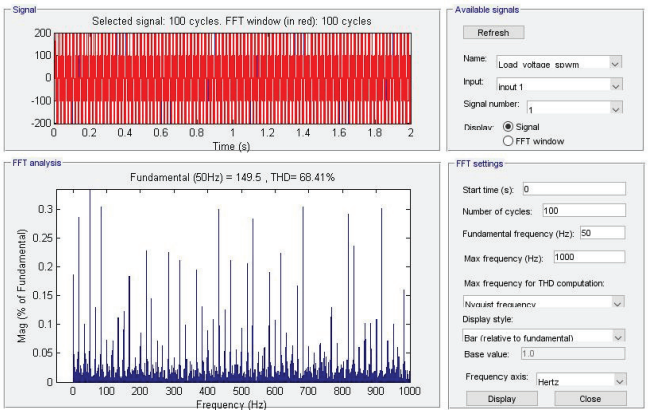


Fig. 9. THD for SPWM using Sawtooth Carrier Wave

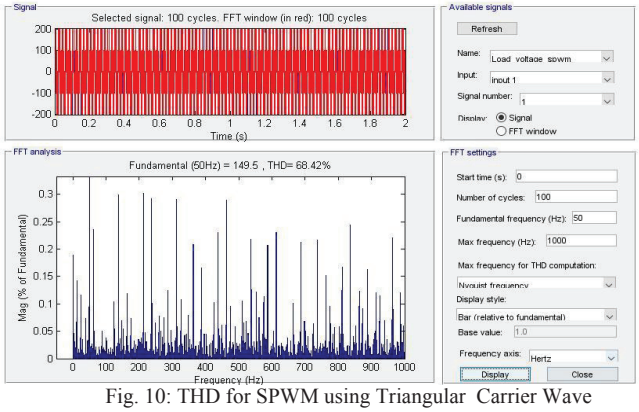


Fig. 10: THD for SPWM using Triangular Carrier Wave

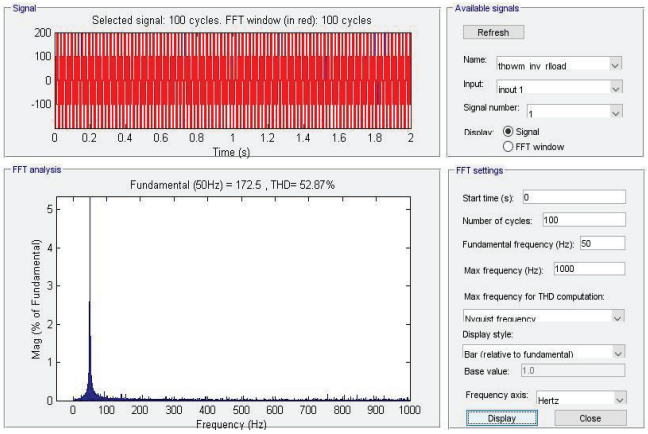


Fig. 11. THD for THIPWM using Sawtooth Carrier Wave

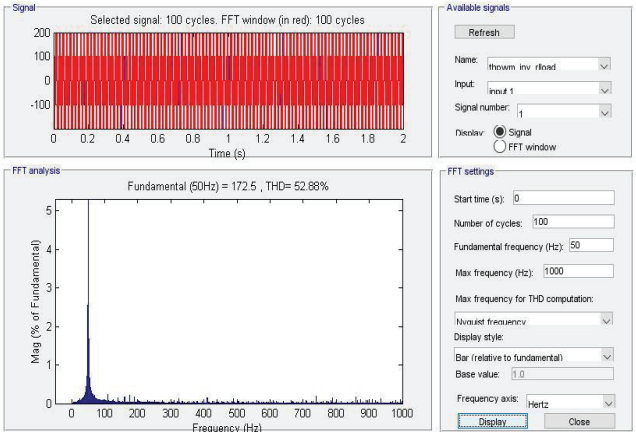


Fig. 12. THD for THIPWM using Triangular Carrier Wave

Figures (9-12) illustrate the various values of Total Harmonic Distortion achieved for different kind of PWM schemes.

TABLE II. COMPARISON OF VARIOUS THD

Carrier Wave	*Reference Wave*	*PWM Technique*	*%THD Value*
Saw-tooth	3 Phase Sinosoidal Wave	SPWM	68.41
Triangular	3 Phase Sinosoidal Wave	SPWM	68.42
Saw-tooth	Modified Sinosoidal Wave	THIPWM	52.87
Triangular	Modified Sinosoidal Wave	THIPWM	52.88

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:59:32 UTC from IEEE Xplore. Restrictions apply.

V. CONCLUSION

This article compares three-phase THIPWM fed inverters with three-phase SPWM fed inverters for different carrier waves. The performance of system generator models for THIPWM and SPWM fed three phase inverter designed in MATLAB-Simulink is sucessfully evaluated.FFT harmonic analysis ratifies the claim of THIPWM being better at DC bus utilization than SPWM.For actual resemblance of the controller design the XSG model is made with the fundamental clock freuency of the FPGA.

REFERENCES

[1] Maiti, A., Choudhuri, S., Bera, J., et al.: 'Development of microcontroller basedsingle phase SPWM inverter with remote control facility'. Proc. Joint Int. Conf.Power Electronics, Drives Energy Systtem/Power India, 2010, pp. 1–5.

[2] M. Kumar and R. Gupta, "Sampled time domain analysis of digital pulse width modulation for feedback controlled converters," IET Circuits, Devices & Systems, vol. 10, no. 6, pp. 481–491, Nov. 2016, doi: 10.1049/iet-cds.2016.0132.

[3] Mayank Kumar, Rajesh Gupta, Time-domain characterisation of multicarrier-based digital SPWM of multilevel VSI, IET Power Electronics, 10.1049/iet-pel.2017.0078, 11, 1, (100-109), (2018).

[4] A. Belkheiri, S. Aoughellanet, M. Belkheiri and A. Rabhi, "FPGA based control of a PWM inverter by the third harmonic injection technique for maximizing DC bus utilization," 2015 3rd International Conference on Control, Engineering & Information Technology (CEIT), 2015, pp. 1-7, doi: 10.1109/CEIT.2015.7233028. .

[5] S. K. Mishra and A. Patel, "Wells turbine modeling and PI control scheme for OWC plant using Xilinx system generator," 2017 4th International Conference on Power, Control & Embedded Systems (ICPCES), 2017, pp. 1-6, doi: 10.1109/spices.2017.8117639.

[6] A.Belkheiri, M.Belkheiri, S.Aoughellanete and A.Rabhi. "FPGA implementation of configurable three-Phase SPWM module ", IEEE Conference Publications, DOI:10.1109/CCCA.2012.6417879. 2nd International Conference on Communications, Computing and Control Application (CCCA), pp. 1-5, DEC 2012 .

[7] Vikas Guptal, Kavita Khare& R. P. Singh, " Efficient Design and Fpga Implementation of Digital Controller Using Xilinx SysGen®", International Journal of Electronics Engineering, 2(1), 2010.

[8] E.Monmasson and M.N.Cirstea"FPGA Design Methodology for Industrial Control SystemsA Review,"IEEE Trans. on Industrial Electronicsvol.54(4),Aug 2007 .

[9] Franklin, G.F., 1.0. Powell and M.L., Workman 1990."Digital Control of Dynamic System". Regarding. MAAddison-Wesley Publishing Company.Yan Yan, Master Thesis KTH/ ICT / ECS - 2007 -88.

[10] J.Jose,G.N.Goyal, M.V.Aware,"Improved Inverter Utilization Using Third Harmonic Injection,"in Proc. PEDES,New Delhi,India ,2010.

[11] M.H. Rashid, Power Electronics Circuits, Devices and Applications, Delhi, India: 3rd ed. Pearson Education, 2006.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:59:32 UTC from IEEE Xplore. Restrictions apply.

Frequency analysis of multi-source single area power system using WCA algorithm

Kirtiman Godara
Department Of Electrical Engineering
Delhi Technological University
Delhi India
godarakirtiman@gmail.com

Narendra Kumar
Department Of Electrical Engineering
Delhi Technological University
Delhi India
dnk_1963@yahoo.com

2021 4th International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE) | 978-1-6654-1429-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/RDCAPE52977.2021.9633690

***Abstract*—— Automatic generation control (AGC) is becoming an essential part of complex power system in order to give good quality power to consumers. When there is load fluctuation happening in power system, it leads to frequency unbalance in power system and due to this complex power system become unstable. To avoid this kind of problem AGC and load frequency control method is used for balancing the power system frequency. In this paper a multi-source single area power system has been presented and is studied using the statistical analysis. In this work proportional-integral (PI) controller has been used as area controller. In order to optimize controller, WCA algorithm has been used. Load frequency control system was simulated on MATLAB with the load disturbances of 0.01p.u.in all the areas. In this work multi-source single area power system LFC has been used for simulation purpose.**

Keywords— Load Frequency Control, Power Systems, water cycle algorithm, proportional control.

I. INTRODUCTION

As the development is taking place through-out the society electricity is now considered as basic need of humanity. Due to wide application and easy convertibility fossil fuels were very much popular. But these fossil fuels are the major contributor to the greenhouse gas emission. Thus its associated negative impact is now causing to shift to non-conventional energy sources like solar, hydro and wind etc. Whereas the availability of non-conventional sources are limited by geographical location of places. Such as solar is limited to between both of the subtropics, where as other area are not able to produce as compare to these. As due to distance of geographical constraints sources are spare less distributed. By the means of transmission line and power cable are interconnected [1].

The interconnection of sources and loads and various facilitating apparatus is called as power system. The limiting of generating stations/sources leads to coupling of several of generators. Sources are being connected to load centers through as mesh of transmission line. As these generators have economically restraints because of which generators are coupled through tie –lines. Tie –lines provide support between the areas for energy exchange during abnormal condition. Tie –lines are used for contractual energy exchange between different areas[2].

The active and reactive power is varying with respect to load (rising and falling trend).The active and reactive power must vary continuously to match the load because whenever there is a mismatch between demand and generation power it causes deviation in system frequency. System frequency is affected by active power of system, therefore as load varies throughout the day so maintaining the system frequency at its nominal value by the governor it's very difficult task. With the help of suitable control strategy the governor valve which controls the real power output of alternator can be regulated automatically. Supplementary control action is required to restore the system frequency at rated value. So primary function of load frequency control (LFC) is to maintain the frequency at nominal value and also maintain the interchange of power between the areas by adjusting the output of generators. Basically load frequency control deals with regulation of real power output of generator and its frequency an incessant progress of AGC strategies over a period of more than five decades. Apart From conventional approach, recently new methods are being implemented in the design and analysis of AGC schemes. Flywheel governor was used earlier to control the frequency of power system but later on a supplementary control was incorporated with the help of signal directly proportional to frequency plus it's integral or integral of summation of deviation in frequency plus tie-line known as area control error (ACE). The secondary control action is slower compared to primary action. Secondary control action keeps the good regulation of frequency[3]. Furthermore, it maintains appropriate real power between the areas. Initial works in automatic generation control were on tie-line based approach [4], [5]. The study of AGC by using an optimization technique based on nonlinear programming might be a maiden attempt. In present research to obtain the results of load frequency operation (LFC) a multiple combination of reheat based thermal, hydro and gas sources are deliberated to keep an eye on LFC complications. Modern control theory and its application by Fosha and Elgerd[6] in power system. In this paper a unified technique dependent for LFC of power system is discussed.

The design of linear PI regulator for two and three area power system was presented by calovic[7]. Two area multi source system was proposed by Kothari[8]. To improve the performance of automatic generation control system under changeable system parameters and operating conditions, various controllers have offered remarkable work over the last decades intelligent technique has emerged in a big way to deal with system design and implementation of control techniques in large power system. various techniques like genetic algorithm (GA)[9], particle swarm optimization (PSO)[10], artificial bee colony (ABC)algorithm[11],bacterial foraging algorithm (BFOA)[12], the controllers are implemented to find out the appropriate instructions to acquire the state with less fluctuations.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:02:19 UTC from IEEE Xplore. Restrictions apply.

In this work a multi-source single area network has been designed in order to mitigate the effect of load variation on power frequency, proportional integral controller (PI) have been used as area controller. Optimization algorithm such as water cycle algorithm (WCA) has been used to tune the parameter of PI controller in a single area multi-source power system network. The water cycle algorithm (WCA) is tested and implemented in the steady state for load disturbance of 1%. Parameters like peak overshoots, peak undershoots, mean, RMS, maximum value and minimum values have been calculated.

In this paper PI controller information is given in section II. Plant description is given in section III. In this section information is given about the governor transfer function and turbine transfer function. Result and discussion is given in section IV. Conclusion is given in section V.

II. PI CONTROLLER

Practically in most of the cases to achieve desired output in a process industry a simple PI controller is used. A PI controller takes the advantages by using both P and I control actions[13]. Major advantage which PI controller provides is by reducing the steady state error because of this reason it is one of the most popular controller which is widely used in industries. PI controller currently used as main control tool due to its simplicity[14]. PI controller always gives desired output due to its effectiveness. The proportional integral gives an output having two terms one is directly proportional and other directly proportional error signal integral. WCA based PI controller has been shown as Fig 2.

Mathematical equation of the PI controller can be given by (1).

$$G_C(s) = K_p + \frac{K_i}{S} \qquad (1)$$

III. PLANT DESCRIPTION

In this paper a multi-source single area power system is considered for load frequency control. In this work a LFC model with three multi-Source name as reheat thermal plant, hydro plant and gas turbine with mechanical hydraulic governor have been used. Basically multi-source single area with its different transfer function are modelled.

A. Plant Model

Multi-source single area consists of governor, reheat turbine, generator, and speed governing system. Transfer function is found with the help of ratio of Laplace transform.

The Laplace transfer function of governor output is defined as –

$$\Delta P_G = \Delta P_{ref}(s) - \frac{1}{R}\Delta F(s) \qquad (2)$$

The Laplace transfer function of governor is as follows:

$$G_G(s) = \frac{\Delta P_V(s)}{\Delta P_G(s)} = \frac{1}{1 + T_G(s)} \qquad (3)$$

Laplace function of the reheated turbine is given by Eq. (4)

$$G_T(s) = \frac{\Delta P_T(s)}{\Delta P_V(s)} = \frac{T_T(s).c + 1}{T_T(s) + 1} \qquad (4)$$

In generator, variation within the increments in power demand and generated turbine power assists as input and the output signal is formed by using the following transfer function

$$\mathrm{G_p}(s) = \frac{\mathrm{K_{ps}}}{\mathrm{T_p}\, s + 1} \qquad (5)$$

SIMULINK model of single area multi source network is designed and is shown in figure 1

B. Objective Function

During the design of PI controller, selection of objective plays a major role. Use of objective function is to optimize the controller based on its performance criteria. In this paper integral time absolute error (ITAE) objective function is used. Mathematical equation of integral time absolute error is given by –

$$ITAE = \int_0^t (|e_i| + \Delta F).t.dt \qquad (6)$$

C. Otimization Technique

The Water cycle algorithm (WCA) is inspired from nature. This algorithm is based on the principle of water cycle process and how river and stream flow towards the sea[16][17].Flow chart of WCA is shown in Fig III.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:02:19 UTC from IEEE Xplore. Restrictions apply.

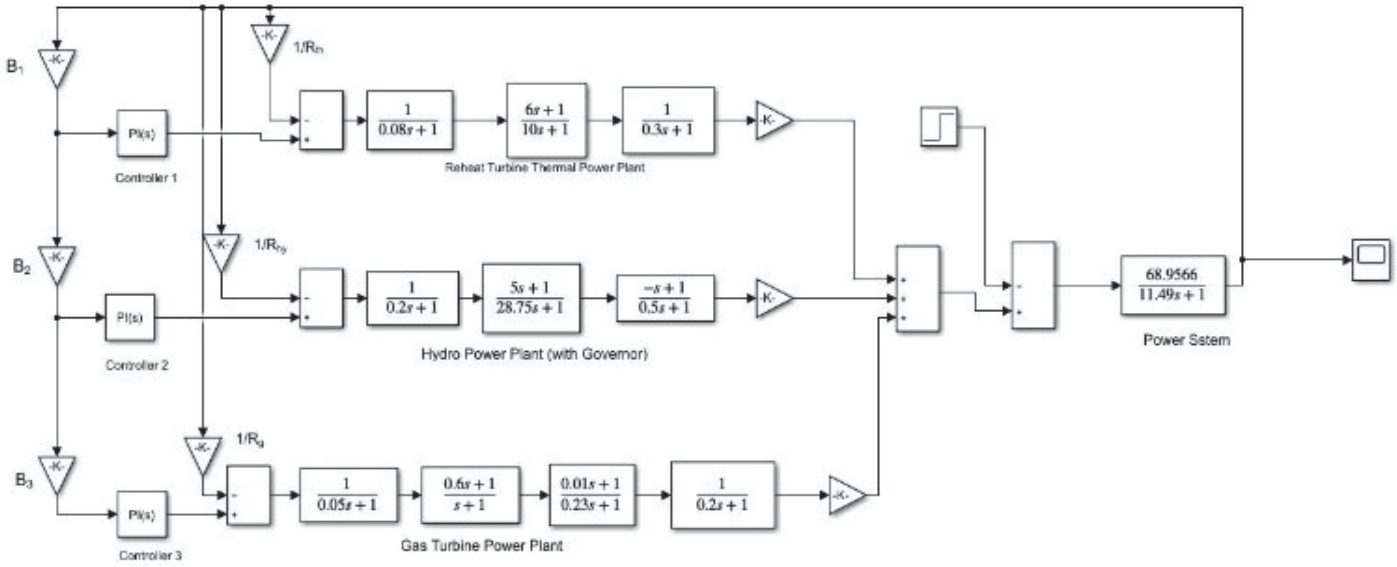


Fig. 1. SIMULINK Model of multisource single area power system [15]

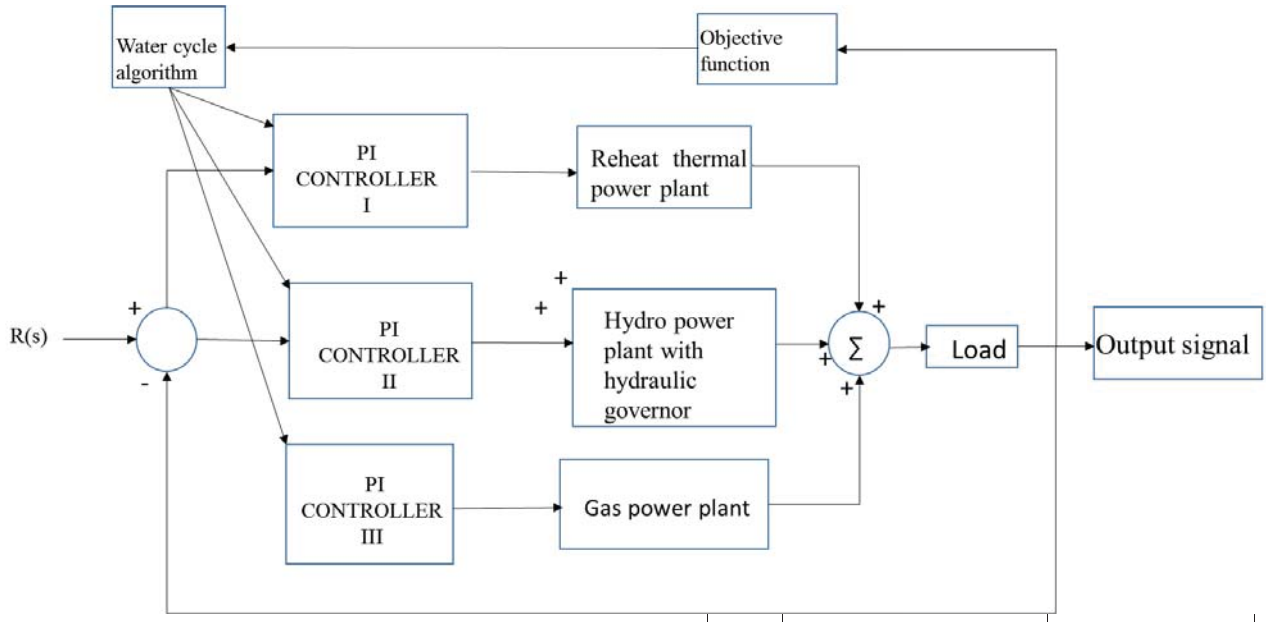


Fig. 2. WCA based PI Controller

IV. RESULTS AND DISCUSSIONS

In this section of paper results obtained from simulation of multi-source single area (reheat thermal, hydro and gas power plant with mechanical hydraulic governor) has been presented. The parameter and their value has been taken for WCA and algorithm. Table I shows the parameters value.

TABLE I. OPTIMIZATION PARAMETERS

Sr. No.	Optimization Parameter	Value
1	Population Size	50
2	Iterations	50
3	Lower bound	-2 for Gain Values 0 for integrator order
4	Upper bound	2 for Gain Values 2 for integrator order

Population size and iteration is taken same for both the algorithms is 50. The objective function is used to optimize the value of PI controller parameter $K_{p1}, K_{i1}, K_{p2}, K_{i2}, K_{p3}, K_{i3}$ for load frequency system. The upper bound and lower bound values were taken +2 to -2 for gain parameters. The results of WCA is presented to show the effectiveness of WCA algorithm. Table II shows the statistical analysis of WCA

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:02:19 UTC from IEEE Xplore. Restrictions apply.

algorithm with PI controller. Which shows the WCA gives better performance.

TABLE II. STATISTICAL ANALYSIS

S. No.	Statistical analysis	WCA
1.	Max value	6.850×10^{-4}
2	Min value	-3.405×10^{-2}
3	Peak to peak value	3.474×10^{-2}
4	Mean	-1.212×10^{-3}
5	RMS	5.456×10^{-3}

Table III represents the optimal value of PI controller using WCA algorithms along with minimize value of ITAE function using algorithm. From the table its clearly shows that WCA performance is better.

TABLE III. PI CONTROLLER PARAMETERS VALUE

Parameter name	WCA
K_{p1}	-1
K_{i1}	-1
K_{p2}	-0.2799
K_{i2}	0.8064
K_{p3}	-1
K_{i3}	-1
ITAE	0.3863

Multi-source single area simulated with PI controller using WCA algorithm and using this algorithm optimized value of controller is obtained. Fig. 2 shows the frequency deviation in area with disturbance is 1%.optimized PI controller used for obtain the frequency deviation of this multi-source single area power system. Fig. 4 shows the frequency deviation in area with load disturbances of 0.01p.u.

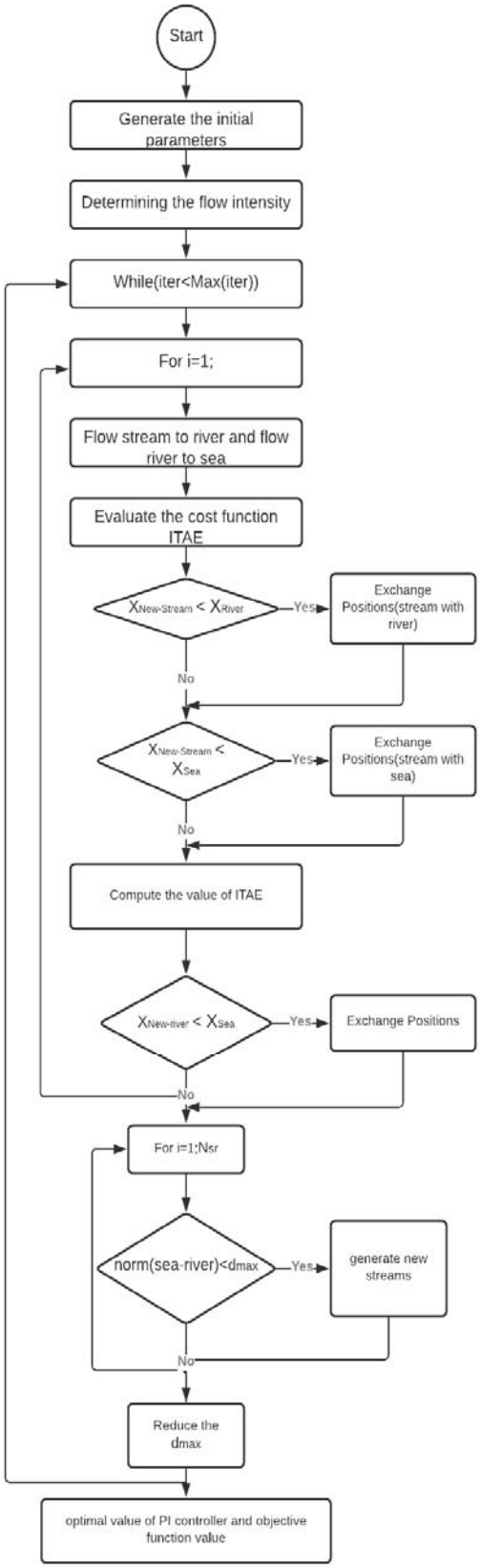


Fig. 3. Flow chart of WCA

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:02:19 UTC from IEEE Xplore. Restrictions apply.

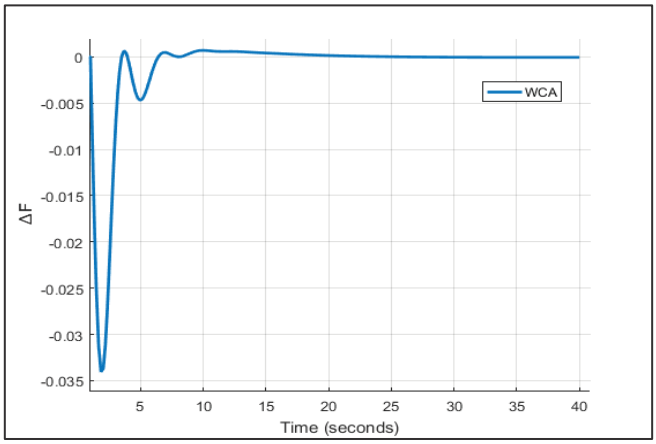


Fig. 4. Frequency Deviation Of area

V. CONCLUSION AND FUTURE SCOPE

Multi source (reheat thermal power plant, hydro power plant and gas power plant) single area has been simulated. In this work PI controller is used as area controller. WCA has been used to tune the parameter of controller. Statistical analysis of PI controller using WCA algorithm for optimization has been done. Some parameters like maximum value, minimum value, peak to peak value mean value, RMS values have been calculated and the frequency behaviour of system when load disturbance happen has also been observed. For future work compare this result with another optimization algorithm in terms of parameters response value.

REFERENCES

[1] "Hadi Saadat - Power Systems Analysis. 3nd Edition-PSA (2011).pdf." .

[2] K. Prabha, "[Prabha Kundur] Power System Stability and Control.Pdf." p. 1176, 1994.

[3] K. P. S. Parmar, S. Majhi, and D. P. Kothari, "Load frequency control of a realistic power system with multi-source power generation," Int. J. Electr. Power Energy Syst., vol. 42, no. 1, pp. 426–433, 2012, doi: 10.1016/j.ijepes.2012.04.040.

[4] N. Cohn, "Some Aspects of Tie-Line Bias Control on Interconnected Power Systems," Trans. Am. Inst. Electr. Eng. Part III Power Appar. Syst., vol. 75, no. 3, pp. 1415–1436, 1956, doi: 10.1109/AIEEPAS.1956.4499454.

[5] C. Concordia and L. K. Kirchmayer, "Tie-Line Power and Frequency Control of Electric Power Systems—Part II," Trans. Am. Inst. Electr. Eng. Part III Power Appar. Syst., vol. 73, no. 1, pp. 133–146, 1954, doi: 10.1109/AIEEPAS.1954.4498803.

[6] M. Nasir, A. Sadollah, Y. H. Choi, and J. H. Kim, A comprehensive review on water cycle algorithm and its applications, vol. 32, no. 23. Springer London, 2020.

[7] M. Ćalović, "Linear regulator design for a load and frequency control," IEEE Trans. Power Appar. Syst., vol. PAS-91, no. 6, pp. 2271–2285, 1972, doi: 10.1109/TPAS.1972.293383.

[8] M. L. Kothari and J. Nanda, "Application of Optimal Control Strategy To Automatic Generation Control of a Hydrothermal System.," IEE Proc. D Control Theory Appl., vol. 135-D, no. 4, pp. 268–274, 1988, doi: 10.1049/ip-d.1988.0037.

[9] K. R. M. Vijaya Chandrakala, S. Balamurugan, and K. Sankaranarayanan, "Variable structure fuzzy gain scheduling based load frequency controller for multi source multi area hydro thermal system," Int. J. Electr. Power Energy Syst., vol. 53, no. 1, pp. 375–381, 2013, doi: 10.1016/j.ijepes.2013.05.009.

[10] V. Jeyalakshmi and P. Subburaj, "PSO-scaled fuzzy logic to load frequency control in hydrothermal power system," Soft Comput., vol. 20, no. 7, pp. 2577–2594, 2016, doi: 10.1007/s00500-015-1659-8.

[11] H. Gozde, M. Cengiz Taplamacioglu, and I. Kocaarslan, "Comparative performance analysis of Artificial Bee Colony algorithm in automatic generation control for interconnected reheat thermal power system," Int. J. Electr. Power Energy Syst., vol. 42, no. 1, pp. 167–178, 2012, doi: 10.1016/j.ijepes.2012.03.039.

[12] L. C. Saikia, N. Sinha, and J. Nanda, "Maiden application of bacterial foraging based fuzzy IDD controller in AGC of a multi-area hydrothermal system," Int. J. Electr. Power Energy Syst., vol. 45, no. 1, pp. 98–106, 2013, doi: 10.1016/j.ijepes.2012.08.052.

[13] B. Popadic, B. Dumnic, D. Milicevic, V. Katic, and Z. Corba, "Tuning methods for PI controller - Comparison on a highly modular drive," IYCE 2013 - 4th Int. Youth Conf. Energy, 2013, doi: 10.1109/IYCE.2013.6604158.

[14] Y. Arya and N. Kumar, "Design and analysis of BFOA-optimized fuzzy PI/PID controller for AGC of multi-area traditional/restructured electrical power systems," Soft Comput., vol. 21, no. 21, pp. 6435–6452, 2017, doi: 10.1007/s00500-016-2202-2.

[15] N. Paliwal, L. Srivastava, and M. Pandit, "Application of grey wolf optimization algorithm for load frequency control in multi-source single area power system," Evol. Intell., no. 0123456789, 2020, doi: 10.1007/s12065-020-00530-5.

[16] M. Barakat, A. Donkol, H. F. A. Hamed, and G. M. Salama, "Controller parameters tuning of water cycle algorithm and its application to load frequency control of multi-area power systems using TD-TI cascade control," Evol. Syst., 2021, doi: 10.1007/s12530-020-09363-0.

[17] M. A. El-Hameed and A. A. El-Fergany, "Water cycle algorithm-based load frequency controller for interconnected power systems comprising non-linearity," IET Gener. Transm. Distrib., vol. 10, no. 15, pp. 3950–3961, 2016, doi: 10.1049/iet-gtd.2016.0699.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:02:19 UTC from IEEE Xplore. Restrictions apply.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Gene expression based inference of drug resistance in cancer

Smriti Chawla[1], Anja Rockstroh[2], Melanie Lehman[2,3], Ellca Rather[2], Atishay Jain[4], Anuneet Anand[4], Apoorva Gupta[5], Namrata Bhattacharya[2,4], Sarita Poonia[1], Priyadarshini Rai[1], Nirjhar Das[7], Angshul Majumdar[4,6,8], Jayadeva[7], Gaurav Ahuja[1], Brett G. Hollier[2], Colleen C. Nelson[2*], Debarka Sengupta[1,2,4,6*]

1. Department of Computational Biology, Indraprastha Institute of Information Technology-Delhi (IIIT-Delhi), Okhla, Phase III, New Delhi-110020, India.
2. Australian Prostate Cancer Research Centre-Queensland, Faculty of Health, School of Biomedical Sciences, Centre for Genomics and Personalised Health, Queensland University of Technology, Translational Research Institute, Brisbane, Australia.
3. Vancouver Prostate Centre, Department of Urologic Sciences, University of British Columbia, Vancouver, Canada.
4. Department of Computer Science and Engineering, Indraprastha Institute of Information Technology-Delhi (IIIT-Delhi), Okhla, Phase III, New Delhi-110020, India.
5. Department of Biotechnology, Delhi Technological University, Shahbad Daulatpur, Main Bawana Road, Delhi-110042, India.
6. Centre for Artificial Intelligence, Indraprastha Institute of Information Technology-Delhi (IIIT-Delhi), Okhla, Phase III, New Delhi-110020, India.
7. Department of Electrical Engineering, Indian Institute of Technology Delhi, Hauz Khas, Delhi, 110016, India
8. Department of Electronics & Communications Engineering, Indraprastha Institute of Information Technology-Delhi (IIIT-Delhi), Okhla, Phase III, New Delhi-110020, India.
9. Genome Institute of Singapore, Agency for Science Technology and Research, Singapore, Singapore.

**Corresponding authors: {debarka@iiitd.ac.in, colleen.nelson@qut.edu.au}.*

Abstract

Inter and intra-tumoral heterogeneity are major stumbling blocks in the treatment of cancer and are responsible for imparting differential drug responses in cancer patients. Recently, the availability of large-scale drug screening datasets has provided an opportunity for predicting appropriate patient-tailored therapies by employing machine learning approaches. In this study, we report a predictive modeling approach to infer treatment response in cancers using gene expression data. In particular, we demonstrate the benefits of considering integrated chemogenomics approach, utilizing the molecular drug descriptors and pathway activity information as opposed to gene expression levels. We performed extensive validation of our approach on tissue-derived single-cell and bulk expression data. Further, we constructed several prostate cancer cell lines and xenografts, exposed to differential treatment conditions to assess the predictability of the outcomes. Our approach was further assessed on pan-cancer RNA-sequencing data from The Cancer Genome Atlas (TCGA) archives, as well as an independent

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

clinical trial study describing the treatment journey of three melanoma patients. To summarise, we benchmarked the proposed approach on cancer RNA-seq data, obtained from cell lines, xenografts, as well as humans. We concluded that pathway-activity patterns in cancer cells are reasonably indicative of drug resistance, and therefore can be leveraged in personalized treatment recommendations.

Introduction

Cancer is a highly complex disease exhibiting varying degrees of genetic and phenotypic heterogeneity within individuals. Despite the apparent overall improvement in prognosis, response to existing therapies (chemotherapies and immunotherapies) are often unpredictable. This is primarily attributable to the clonal diversity of cancer cells and associated phenotypically altered non-malignant cells in the tumor microenvironment. These pose a substantial hindrance to the optimal therapeutic management of the disease[1,2]. As a rescue, presently, therapies that work by targeting specific molecules or mutations are highly advocated. EGFR expression & mutations, KRAS mutations, BCR-ABL fusions, and HER2 overexpression are examples of popular therapeutic targets in cancer[3]. Such therapies are frequently rendered futile or counterproductive due to the unavailability of known molecular targets in patient tumors. As such, it's concluded that drug targets or mutational status alone are incomprehensive as indicators for targeted therapies[4,5]. Furthermore, administering a targeted therapy without considering drug resistance as a consequence may worsen patient survival. As such, early inference of drug response, based on pretreatment molecular portraits of cancer has become a necessity[6,2].

In recent years, the availability of large-scale pharmacogenomic databases has propelled predictive personalized oncology research[2]. Cancer Cell Line Encyclopedia (CCLE)[7] and Genomics of Drug Sensitivity in Cancer (GDSC)[8] are noteworthy among these[9]. These studies embody an expansive knowledgebase comprising high-throughput screening studies entailing more than 1000 cell lines with hundreds of anticancer drugs[2]. Concurrently, The Cancer Genome Atlas (TCGA)[10] serves as another rich database featuring gene expression profiles of primary tumors

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

spanning multiple cancer types, with associated clinical meta-data and drug-response annotations. This wealth of data has enabled the modeling of drug response based on molecular profiles. Various machine learning methods have been proposed for drug response prediction in cancer. Jia et al. proposed a deep variational autoencoder for imputing drug response through compression of multiple genes into latent vectors in low dimensional space[9]. Ammad-Ud-Din, Muhammad, et al.[11] reported a kernelized bayesian matrix factorization based drug response prediction through incorporating prior knowledge of pathway-drug associations. Another approach utilized matrix factorization with similarity regularization for drug response prediction in cell lines by employing chemical structures of drugs and gene expression profiles[12]. Chang and colleagues proposed CDRscan, a convolutional neural net that leverages mutational signatures for predicting drug effectiveness[13]. Sakellaropoulos, Theodore, et al. reported gene expression data based deep neural network for drug response prediction that outperforms ElasticNet and Random Forest[14].

By carefully surveying these methods, we identified two critical scopes for improvement. First, the majority of the past studies do not consider structural properties of drugs as features (explanatory variables in predictive tasks). As a result, the machine learning models learn suboptimally and fail to make predictions on drugs that are not part of the training data. Secondly, gene expression levels are considered as independent variables, ignoring their pathway-specific combinatorial implication. Past studies have demonstrated the utility of using pathway enrichment scores for various downstream analyses, as opposed to gene expression values[15,16]. Notably, in our previous works, we have shown how pathway projection facilitates better modeling of biological processes[16,17]. As an added advantage, data integration based on pathway enrichment scores subsides batch-effects[18]. While single-cell RNA-seq (scRNA-seq) enables characterization of cellular heterogeneity in tumors, there's very little visible effort to leverage this fine-grained molecular information to predict drug-response at sub-clonal resolution. This is primarily due to the fact that most available training data, as indicated above, are bulk expression profiles, and training on bulk RNA-seq and testing on scRNA-seq is expected to give rise to misleading predictions. Pathway projections of scRNA-seq/bulk RNA-seq profiles reasonably alleviate this problem. Notable in this regard is the work by Suphavilai, Chayaporn, et

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

al.,[19] that describes a drug response prediction approach in head and neck cancer, leveraging scRNA-seq profiles. The authors, however, did not explore the utility of drug descriptors to generalise the prediction model.

In this work, we used various machine learning and deep learning approaches (DNN) to train models for the prediction of drug-response in both *in vitro* and *in vivo* settings. We made use of publicly available high-throughput screening data (CCLE and GDSC) concerning a multitude of drug-cancer cell line pairs, as well as patient data from TCGA. We validated the cell line based model on multiple scRNA-seq datasets of established cancer cell lines.

Convinced by the reproducibility and overall performance of the cell-line model, we delved deeper into similar prediction tasks pertaining to in-house prostate cancer cell lines and animal models under differential treatment conditions. Prostate cancer is the most common malignancy in men, yet for patients presenting with advanced cancer, treatment options are limited. Androgen deprivation is an effective therapeutic strategy widely used in clinical practice. It exploits the unique dependence of prostate cancer on androgen signaling for growth and progression. Although androgen deprivation therapy is effective in most patients, the effect is only palliative and cancer cells eventually become resistant with the emergence of metastatic castration-resistant prostate cancer (CRPC). To date, only a limited number of agents have shown effectiveness and have been approved for the treatment of CRPC, yet their conferred survival benefit is only minimal, and acquired resistance to the drugs eventually emerges. Thus, appropriate drug selection and combination remain crucial in the dynamically cancer evolving landscape to derive maximum benefit for the patients[20,21,22]. We employed our internal bulk RNA-seq profiles of prostate cancer cell lines under different treatment conditions for further validation. To ensure the cross-sample utility of our model, we examined our LNCaP xenograft dataset mirroring *in vitro* treatments. For this, we utilized LNCaP xenografts from a prostate cancer progression study where the tumors were harvested at different stages, namely Pre-castration (PRE-CX), post castration (POST-CX), castration-resistant prostate cancer (CRPC), and while on enzalutamide (ENZ) treatment during the course of tumor progression. Our results revealed clinically and biologically relevant associations of drugs and pathways in terms of

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

resistance and sensitivity. Along the same lines, we benchmarked the efficiency of the model, trained on patient samples from TCGA on RNA-seq profiles of pre-treatment, and matched post relapse drug-resistant BRAF mutant melanoma cancer patients. Our study connects a systematic drug response prediction pipeline with layered *in vitro* and *in vivo* validations involving cell lines, xenografts, and patients, which is the most important prerequisite for the clinical implementation of such approaches.

Results

Overview of approach

In this study, we model drug response using gene expression data in both *in vitro* and *in vivo* settings. For this, we leveraged Cancer Cell Line Encyclopedia (CCLE) cancer cell line gene expression RNA-seq bulk profiles. We formulated the prediction of drug response in terms of Z-score (LN IC50) as a regression task and constructed a suitable deep neural network (DNN) architecture. We performed rigorous assessment of drug response prediction in terms of LN IC50 and Z-score (LN IC50) on several datasets. Previous studies have reported use of IC50 for modelling drug response[23]. However, our analysis didn't reveal biologically reasonable results with LN IC50 (data not shown). Hence, we resorted to Z-scores (LN IC50) for our analysis. First, we converted bulk RNA-seq gene expression profiles of cell lines from the CCLE project into pathway scores using GSVA[15]. Second, we integrated the cell line pathway scores matrix with the simplified molecular-input line-entry system (SMILES) embeddings (numerical descriptors of molecules) of the drugs. We sourced drug response information for the CCLE cell lines from the Genomics of Drug Sensitivity in Cancer (GDSC) database. The summary of the datasets used is provided in **Supplementary Fig.1.** We found 550 cell lines common in both the databases screened against 173 unique molecular compounds for which SMILES notations were available. We retrieved SMILES of 173 compounds using PubChemPy[24] and converted them into vector embeddings of size 100 using the SMILES2Vec python package[25]. The final matrix used for training the model constituted 80056 cell line-drug combinations in rows and 1429 features entailing 1329 pathway score vectors and vectors of size 100 each for molecular descriptors in columns.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

These vectors form a set of explanatory variables to predict drug response as obtained from the screening experiment datasets **(Fig. 1a)**. We used the open-source software Keras[26] to build a DNN architecture **(Fig. 1b)**. We evaluated the performance of the DNN model on the CCLE bulk RNA-seq cell lines test dataset and achieved a coefficient of determination (R^2) = 0.60; Pearson's correlation coefficient = 0.78; *P*-value p < 2.2e−16 **(Fig. 1c)**. We then compared the performance of the DNN model with the other two commonly used machine learning models: random forest (RF) and elastic net that have been used by previous studies for drug response prediction[27,28,29] **(Supplementary Fig. 2).** The DNN architecture outperformed the other two models. As a control, we also evaluated the performance of these three models using gene expression vectors for 500 genes (as opposed to pathway score matrix) selected based on the coefficient of variation square (CV^2). We observed that in gene-based models DNN model is better in comparison to RF and elastic net **(Supplementary Fig. 3a, b)**. Further, our results revealed that the performance of pathway-based models is better than gene-based models in the case of the elastic net and DNN. These findings substantiate our choice of pathway scores over gene expression values. To ensure an unbiased evaluation of our DNN model, we applied our CCLE bulk RNA-seq model to external datasets including both bulk RNA-seq and scRNA-seq profiles. We used the Kinker, G. S. et al. scRNA-seq cell line dataset[30]. After processing, this dataset consisted of 17279 drug-cell line combinations entailing 116 cell lines and 173 drugs **(see methods)**. We were able to achieve a coefficient of determination (R^2) = 0.42; Pearson's correlation coefficient = 0.65; *P*-value p < 2.2e−16 **(Fig. 1d)**. As expected, we did not observe a reasonable correlation in scRNA-seq profiles than we achieved on bulk RNA-seq profiles. Further, we benchmarked our model using another scRNA-seq dataset from a previously published study by Lee et al.*[31]* consisting of treatment-naive breast cancer cells (MDA-MB-231) and a population of cells that had gained sensitivity to paclitaxel. In this study, the metastatic MDA-MB-231 cells were exposed to a paclitaxel drug. Most of the cells died after 5 days of exposure. However, some of the residual cells, cultured in drug free medium after withdrawal of drug, recommenced proliferation and established clones. Interestingly, these cells became more sensitive to paclitaxel on reexposure. To verify that these cells are more sensitive to paclitaxel, we applied the DNN model to this dataset and found that

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

our model predicted the paclitaxel-sensitive population to be more sensitive to paclitaxel in comparison to the treatment-naive population, corroborating with the original findings **(Fig.1e)**.

Evaluation of models trained on CCLE bulk RNA-seq profiles using independent datasets

Despite therapeutics advances in prostate cancer, treatment options remain limited and emergence of treatment resistance poses significant challenges. Therefore, there is an unmet need for selection of optimal drug for prostate cancer treatment[32,33]. We independently validated our DNN model trained on CCLE bulk RNA-seq profiles using our in-house prostate cancer datasets. We applied the DNN model to bulk RNA-seq profiles of five untreated prostate cancer cell lines with two biological repeats of each cell line. We predicted drug responses for each of these ten samples for 155 drugs tested against prostate cancer cell lines in the GDSC database targeting various cellular pathways. Overall, we found the Androgen Receptor (AR) negative prostate cancer cell lines DU145 and PC3 were predicted to be more resistant to these therapeutic drugs, while the AR-positive cell lines LNCaP, DUCAP, and VCAP were more sensitive **(Fig. 2a)**. Of these five cell lines, LNCaP cells were predicted to be most sensitive to these drugs **(Fig. 2b)**. More specifically, the model predicted the sensitivity of LNCaP cells to PI3K/mTOR signaling pathway targeting drugs, in particular AKT inhibitors, afuresertib, and uprosertib **(Supplementary Fig. 4a)**. The LNCaP cells had high GSVA pathway scores for mTOR-related signaling, which may contribute to the predicted sensitivity to these AKT targeting drugs **(Supplementary Fig. 4b)**. When comparing the predicted Z-scores with the corresponding GDSC Z-scores, we observed a high Pearson correlation of 0.78 and 0.84 with a statistically significant *P*-value $p < 2.2e{-}16$ for the two biological replicates of LNCaP cells, respectively **(Fig. 2c)**.

We were further interested in testing how the drug response prediction was altered when LNCaP cells are cultured with androgen receptor (AR) agonist dihydrotestosterone (DHT) as compared to the vehicle control (VEH) in androgen deprived media conditions; and furthermore, how treatment with clinically approved AR antagonists bicalutamide (BIC), enzalutamide (ENZ), and apalutamide (APA) under these conditions affected the predicted sensitivity pattern. We used averaged GSVA scores of the biological replicates for the downstream analysis. Overall, cells

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cultured with DHT were predicted to be more sensitive to therapeutic drugs, while the cells cultured with no DHT and AR antagonists were more resistant **(Fig. 2d, Supplementary Fig. 5a)**. The cells cultured in the presence of DHT had high GSVA pathway scores of proliferation-associated pathways. Androgens are involved in stimulating prostate cancer cell proliferation[34,35]. This supports the notion that actively proliferating cells are more sensitive to particular anti-cancer drugs, while cells in a cytostatic or quiescent state are more resistant (**Fig. 2e)**. Notably, the addition of AR antagonists in the presence of DHT did not fully reverse the predicted DHT conferred drug sensitivity. However, for a small subset of drugs mainly targeting the PI3K/mTOR signaling pathway, AR antagonists decreased the predicted drug sensitivity **(Fig 2f)**. In particular, our DNN model predicted that treatment with ENZ in the presence of DHT compared to DHT alone rendered the cells more resistant to AKT inhibitors, in particular uprosertib and afuresertib **(Fig. 2d, f)**. In the absence of androgens, ENZ and BIC treatment predicts increased resistance to drugs targeting EGFR signaling, with gefitinib and afatinib having the most profound effect. Whereas the addition of DHT alone increased drug resistance to EGFR targeting agents irrespective of the inclusion of AR antagonists **(Supplementary Fig. 5d)**.

Next, we evaluated our ability to predict drug response in LNCaP derived xenografts. Xenografts have emerged as useful *in vivo* cancer models for directly investigating therapeutic response and predicting anti-cancer drug response in patients with cancer of a similar phenotype. To this end, we used in-house data from LNCaP xenografts derived from a large and well-annotated prostate cancer progression study investigating responsiveness and subsequent resistance to therapies targeting the AR. LNCaP xenograft tumor establishment and initial growth are dependent on androgens in male mice (PRE-CX). Upon castration, AR activity and tumor growth is suppressed (POST-CX), however, this responsiveness to castration reproducibly gives way to castration-resistance (CRPC). Further treatment of CRPC with ENZ initially provides a therapeutic response (ENZ Sensitive; ENZS), however, resistance emerges in time (ENZ Resistant; ENZR) **(Fig. 3a)**. Using our DNN model, we predicted the drug response for each of the 54 samples across this spectrum of successive therapeutic responsive and resistance states. The LNCaP xenograft tumor samples clustered into three main groups based on their overall predicted sensitivity to the 155 drugs in the analysis **(Fig. 3b, Supplementary fig. 6a)**. Cluster 1 samples had the most resistant tumors,

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

which correlated to their lower proliferative index. Cluster 1 consisted of one CRPC sample and the remaining were ENZ-treated tumor samples (7 of the total 15 ENZR and 9 of 11 ENZS samples). In contrast, cluster 3 samples had the highest predicted overall sensitivity to the 155 drugs, which may be attributed to their higher proliferative index, indicated by higher GSVA pathway scores for cell proliferation-associated gene sets **(Fig. 3c, d)**. Cluster 3 consisted of all PRE-CX mice tumors and subsets of the resistance types (5 of 10 POST-CX; 7 of 10 CRPC; and 4 of 15 ENZR samples). ENZR tumors were distributed across all three clusters, suggesting that resistance to ENZ occurs through different underlying mechanisms and more detailed analyses may uncover differential susceptibility to particular drugs in ENZR. This suggestion of multiple ENZ resistance mechanisms was strengthened by the multimodal distribution of predicted Z-scores for ENZR tumors compared to the uniform distribution for ENZS tumors **(Fig 3e)**. In contrast to ENZS, ENZR samples were predicted to develop some level of sensitivity to a subset of drugs. ENZR samples tended to have higher GSVA scores of proliferation-associated pathways relative to ENZS samples however, this did not reach statistical significance **(Fig 3f)**. ENZR tumors were predicted to be more sensitive to EGFR targeting drugs than any other tumor type in the study **(Fig 3g)**, with sapitinib having the most profound effect **(Supplementary Fig S6b)**. While we achieved promising informative results on the drugs within our training set, we could also predict biologically relevant responses for drugs not included in the training set— APA, BIC, and ENZ. We observed sensitivity to AR antagonists in the PRE-CX, POST-CX and CRPC groups. Notably, we found responsive ENZS tumors from mice actively treated with ENZ were predicted to not benefit from additional AR antagonists. **(Fig 3h)**.

Evaluation of models trained using TCGA patient tumor profiles

The Cancer Genome Atlas (TCGA) furnishes a large compendium of datasets spanning multiple cancer types with gene expression profiles of primary patient tumors and clinical response information. The clinical drug response data includes information about patient demographics and patient response for the drug administered. We aimed to perform classification by stratifying patients as responders who had a complete or partial response and non-responder for the patients who had clinically progressive or stable disease. To make the best use of the patient

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

gene expression profiles and clinical drug response data, we used AutoML provided by open-source H2O.ai (https://docs.h2o.ai/h2o/latest-stable/h2o-docs/automl.html)[36] in R to build a classifier. The summary of the dataset used is provided in supplementary data **(Supplementary Fig. 7)**. As discussed before, for CCLE bulk RNA-seq cell lines data, we integrated our processed TCGA pathway score matrix **(see methods)** with the vector embeddings of 139 unique drugs that were sourced from TCGA clinical drug response data. This matrix constituted 3108 patient-drug combinations in rows and 1427 features entailing 1327 pathway score vectors and vector embeddings of drug compounds of size 100 each, forming a set of explanatory variables to predict drug response as obtained from TCGA drug response data. For classification, we split this dataset and used 80% of the dataset for training the AutoML models, and 20% of the dataset was used for testing purposes. Our training dataset comprised of 2486 drug-patient combinations and the testing dataset consisted of 622 drug-patient combinations. AutoML is a platform that enables users to automatically train and tune multiple models by specifying a maximum number of models, thus automating machine learning workflows. We specified 20 maximum models to be trained with five fold cross validation. This resulted in 22 models, of which 20 models spanned across families of models, including GBM, XGBoost, DeepLearning, DRF, XRT models, and two stacked ensemble models. The two 'stacked ensemble models' were automatically trained and correspond to — one based on 'all prior models' that have been trained, while the other one is based on 'each family's best model'[36]. We evaluated the performance of the test dataset on these 22 trained models and found the 'stacked ensemble model', that was trained based on 'all prior models' outperform all others with an AUC=0.96, AUCPR= 0.98, and Gini=0.91 **(Fig. 4a)**. Other evaluation metrics are reported in the supplementary data **(Supplementary Table S1)**. We used the best model for all subsequent prediction tasks. We also reported cross-validation performance metrics for these models **(Supplementary Table S2)**. Additionally, TCGA data have recorded survival information for patients. Using this information, we conducted survival analysis on the test dataset comprising 622 patients by stratifying patients based on the median value of the predicted probability of response. As shown in **Fig. 4b**, patients in group 1 who had a probability of response greater than the median value had better survival outcomes ($P < 0.0001$). We estimated the 5-years survival probability of group 1 as 0.774 and group 2 as 0.349. Then we

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

sought to traverse the relevance of cancer stages in influencing the significant survival for group 1. For this, we visualized the frequency of cancer stages across different cancer types for group 1 and group 2. We observed that for the majority of cancer types, stage information is not known for most patients. Notably, in group 1, BRCA cancer type has the highest number of patients (n=34) belonging to stage iia, however, only 8 patients belonged to this stage in group 2. In our test dataset, we have a total of 42 BRCA individuals in stage iia, with 80.95 % of these patients falling into group 1 and 19.04 % falling into group 2 **(Supplementary Fig. 8)**. Our results are concordant with previous studies that early detection of cancer is associated with improved overall survival[37]. Further, **Supplementary Fig. 9** shows confusion matrices for group 1 and group 2 depicting the concordance of predicted and actual drug responses in patients. Next, we evaluated our model using a dataset for which pre and post-treatment RNA-seq profiles and clinical response information are available[38]. We predicted the drug response of RAF inhibitor (dabrafenib) and MEK inhibitor (trametinib) in three pre-treatment and matched post-lapse BRAF-mutant RAF/MEK inhibitor-resistant melanoma patients. **Fig. 4c** shows the journey of patient 1. The patient was diagnosed with Stage IIIC melanoma. After 1 year of initial treatment, the patient again showed signs of recurrent disease and underwent pre-treatment biopsy and intensity-modulated radiation therapy (IMRT). This patient exhibited both BRAF V600E and V600K mutations and was recruited into a Dabrafenib and Trametinib phase I/II study. After three months, the patient was removed from the study due to the development of resistance. Then the patient was treated with anti-PDL1 antibody for 4 months until progression and subsequently treated with four cycles of Ipilimumab. The patient died of his condition approximately nine months after the stoppage of RAF and MEK inhibitors. The journeys of other patients are documented in supplementary data (**Supplementary note 1)**. Wagle, Nikhil, et al. reported the presence of $MEK2^{Q60P}$ mutation, BRAF Splice Isoform, and BRAF amplification in the post-treatment patient one, patient two, and patient three, respectively, as revealed through whole-exome sequencing (WES) and RNA-seq. These alterations seem to be a potential cause of conferring resistance to RAF/MEK inhibitors in these patients[38]. Notably, our prediction results revealed a similar trend, whereby our probability of response of dabrafenib and trametinib in pre-treatment patient one and patient two is higher than post-treatment **(Fig. 4d, e)**.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Furthermore, we were correctly able to predict these patients as responders. This aligns with the original annotations of the study as these patients are categorized as a partial response based on Response Evaluation Criteria In Solid Tumors (RECIST)[39]. In contrast, patient three was correctly predicted to be resistant to dabrafenib for both pre and post-treatment **(Fig. 4f).** This is concordant with the original study where this patient was categorized as a stable disease based on RECIST criteria.

Discussion

Predicting drug response of cancer cells is of paramount importance for personalized oncology. In this study, we designed a deep neural network (DNN) model to predict the drug response by leveraging pathway activity scores derived from gene expression profiles of cell lines from CCLE and corresponding vector embeddings of drug compounds that were sourced from GDSC. Due to the explicit use of pathway enrichment scores, our model also highlights underpinning biological mechanisms that contribute to drug resistance. Furthermore, our pathway-based prediction approach allowed us to accurately infer cell fates upon treatment from single cell expression data. This can open doors for drug response prediction at sub-clonal resolution using tumour scRNA-seq data.

We evaluated the DNN model predictions by showing a reasonable correlation between the experimental Z-scores (LN IC50) for LNCaP cells used in the training data set and the model-predicted Z-scores. Further, we noted that *PTEN* negative LNCaP cell line was predicted to be sensitive to PI3K/mTOR signaling targeting drugs. Notably, previous studies have broadly suggested that increased PI3K/AKT/mTOR signaling is associated with sensitivity to PI3K inhibitors in LNCaP cell lines and other PTEN null cancer cell lines. Furthermore, mTOR inhibitors might be effective against PTEN null tumors[40]. With our predictive model, we could spot the sensitive phenotype of LNCaP cell lines to PI3K/mTOR inhibitors. This aligns with the previous findings in the LNCaP cell line.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

With the confidence gained from cell line based validations, we further explored how drug sensitivity prediction changes upon drug treatments, drug-responsive and resistant states in LNCaP cells and xenografts. We sought to explore how these findings align with the known biology surrounding these treatments and states. In the presence of androgens, LNCaP cells were predicted to be more sensitive to cancer therapeutics that target highly proliferative cells. This is an expected result since androgens drive proliferation in AR-positive prostate cancer cell lines and xenografts[34,35]. AR antagonists are the primary treatment option for metastatic prostate cancer, which on a molecular level antagonize cellular androgen response pathways. With AR antagonist treatments, LNCaP cells were predicted to have pronounced similarities and differences for ENZ, APA, and BIC treatments, mirroring the complex biology underlying the treatment responses. The strongest reversal by an AR antagonist of DHT conferred sensitivity was observed for drugs targeting the PI3K/mTOR pathway, with ENZ showing the most profound effect. In the absence of androgens in all conditions, LNCaP was predicted to be resistant to drugs targeting DNA replication and genome integrity, namely irinotecan, cisplatin, and talazoparib. While ENZ treatment was uniquely predicted to increase the DHT conferred sensitivity to these drugs, BIC and APA partially decreased the DHT conferred sensitivity **(Supplementary Fig. 5b, c)**. These differential effects are important to be considered for combinatorial therapy in a given context of treatments and treatment resistance with other cancer therapeutics.

To further test the applicability of the drug sensitivity predictions, we have used data derived from an internally generated and biologically well-annotated prostate cancer xenograft model, following the progression from early androgen-responsive to CRPC, and then to ENZ treatment responsive and ENZ resistant states. The sensitivity predictions highlight changing vulnerabilities of the tumors in different stages of progression and treatment. Notably, we observed that ENZR tumors are predicted to develop new susceptibility to novel therapeutics providing a new window of opportunity for new therapeutic strategies. For instance, our model predicted sensitivity to specific drugs targeting the EGFR signaling pathway in case of ENZ treatment. While prostate tumors in an androgen replete setting are predicted to be highly resistant to EGFR targeting drugs, a distinct vulnerability is developed during progression to ENZR. As these drugs targeting the EGFR pathway are approved for cancer treatment in other cancer types, this may

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

facilitate the clinical evaluation of combination therapy in prostate cancer patients treated with ENZ or those that have developed ENZR. Previous studies have suggested that a combination of ENZ and EGFR targeting inhibitors might be an effective therapeutic strategy in overcoming ENZ resistance[41]. Further laboratory and preclinical studies into the molecular mechanism behind our predicted sensitivity pattern are needed to confirm these findings.

Additionally, our AutoML model, trained on TCGA tumor RNA-seq data, revealed substantially improved overall survival in patients, predicted for partial or complete drug response. Further, testing of the TCGA model using the external BRAF-mutant melanoma dataset resulted in clinically relevant predictions for the patient samples. As expected, the probability of response for the patients classified in the partial response category was higher than post-treatment due to acquired resistance to the first line of therapy. This suggests that top drugs predicted by our method, for example, cyclophosphamide, cisplatin, might serve as alternative therapies in combination with other drugs to overcome acquired resistance.

There are some technical limitations of our work. In TCGA clinical drug response data, the same patients have been treated with multiple drugs, so we have treated them individually keeping the transcriptome identical.

Overall, this is among the first works connecting bioinformatic predictions of drug response to clinically explainable observations, both in *in vivo* and *in vitro* settings. Further, to our knowledge this is the first focused study on prostate cancer, investigating the potential of tumor bulk RNA-seq data in predicting drug response.

Methodology

Prostate cancer cell lines and culture

The human prostate cancer cell lines LNCaP, DuCaP, VCaP, DU-145, and PC-3 were cultured in Phenol-red free RPMI medium supplemented with 5% fetal bovine serum (FBS) in a humidified incubator at 37°C and 5% CO2 and were harvested for RNA extraction during their exponential growth phase.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In vitro LNCaP cell line treatments

The human prostate cancer cell line LNCaP (#CRL-1740™ clone FGC, purchased from ATCC) was seeded into Phenol-red free RPMI medium supplemented with 5% fetal bovine serum (FBS) and cultured for 72 hr in a humidified incubator at 37°C and 5% CO2. This was followed by a 48 hr incubation in androgen-depleted conditions using medium + 5% charcoal-stripped serum (CSS). Treatment with the androgen targeting drugs (ATTs) enzalutamide (10uM), bicalutamide (10uM), and apalutamide (10uM) was performed for 48 hr, either in the absence or presence of 10 nM dihydrotestosterone (DHT, dissolved in EtOH).

Establishment of LNCaP xenografts

For the *in vivo* tumor progression study, xenografts were established by subcutaneous injection of one million LNCaP cells into the flank of male NOD-SCID mice. At a tumor size of 200 square mm, mice were either surgically castrated or received mock surgery for the PRE-CX group. Tumors from the PRE-CX group were harvested when the tumor size reached 1000 square mm. Tumors from the POST-CX (post castration) group were harvested 1-week post castration when serum PSA (Prostate-Specific Antigen) reached a nadir. Tumors from the CRPC (Castrate-Resistant Prostate Cancer) group were harvested once the tumor size reached 1000 square mm post castration. For the 'ENZ' groups, daily treatment with 10 mg/kg ENZ commenced as serum PSA began to rise post castration. Tumors were harvested either at PSA nadir (ENZS) while on enzalutamide treatment or when the tumor size had reached 1000 square mm despite enzalutamide treatment (ENZR).

RNA extraction, library preparation, and bulk RNA-sequencing

For mRNAseq, total cellular RNA was extracted using the Norgen RNA Purification PLUS kit #48400 (Norgen Biotek Corp., Thorold, Canada) according to the manufacturer's instructions, including DNase treatment. RNA quality and quantity were determined on an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, USA) and Qubit®. 2.0 Fluorometer (Thermo Fisher Scientific Inc, Waltham, USA). Library preparation and sequencing was done using the 'Illumina

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TruSeq Stranded mRNA Sample Prep Kit (strand-specific, polyA enriched, Illumina, San Diego, USA) with an input of 500 ng - 1 ug total RNA (RIN>8), followed by paired-end sequencing with a read length of 100-150 bp and yielding about 30-60 M read pairs per sample.

RNAseq raw data was processed through a custom-designed pipeline. Raw reads were assessed with FastQC[42], then trimmed using TrimGalore[43], followed by alignments to the human genome (GRCh38 / hg38) and transcriptome (Ensembl.v.99 / Gencode.v.33, Jan-2020) using the STAR[44] aligner and read quantification with RSEM[45]. For xenograft samples, STAR alignment was done against a chimeric human+mouse reference (mouse: GRCm38 / mm10, Gencode.v.M24 / Ensembl.v.99, Jan-2020), followed by RSEM read quantification. TPM values from the RSEM output were used for GSVA scoring.

Gene expression data of cancer cell lines

To predict and investigate anti-cancer drug response measured by Z-score of half-maximal inhibitory concentration (LN IC50), we used publicly available TPM (transcript per million) normalized RNA-seq gene expression profiles of 1019 Cancer Cell Line Encyclopedia (CCLE)[7] cell lines quantified using the RSEM (RNA-Seq by Expectation-Maximization) software. The corresponding drug response information for the cell lines was sourced from the GDSC2 dataset of the Genomics of Drug Sensitivity in Cancer (GDSC) database[8]. In the GDSC2 dataset, some drug-cell line pairs have multiple LN IC50 measurements. In such cases, we averaged the LN IC50 values to avoid ambiguity. Then, LN IC50 values of all the drugs were transformed into Z-scores by using the formula $(x-\mu)/\sigma$. The variable x represents LN IC50 values for a drug across cell lines, whereas μ & σ, the mean and standard deviation respectively. The RNA-seq gene expression profiles of 550 CCLE cell lines overlapping with the GDSC2 dataset cell lines were used for the model training. This matrix contained 57820 Ensembl Gene IDs which were converted into official gene symbols using gencode.v19.genes.v7_model.patched_contigs.gtf annotation file. This resulted in multiple Ensembl gene IDs corresponding to some of the individual gene symbols. We considered the average expression in such cases. At this stage, our expression matrix comprised 54301 genes

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and 550 cell lines. This matrix was subjected to log2 transformation with the addition of a pseudo count of 1.

Tumour RNA-seq data

Analogous to cell lines, on tumour mRNA sequencing data from TCGA, we modeled drug-response in terms of responder and non-responder. TCGA RNA-seq data was downloaded from the Broad GDAC firehose[46] encompassing 33 tumor types. We used Illumina HiSeq RNA-seq v2 data processed at the gene level using RSEM[45]. The clinical drug response information for the patient samples was fetched from the NCI Genomic Data Commons portal[47]. Drug names and response information were collected from clinical metadata and rectified for typographical and spelling errors and to harmonize commercial names and molecular drug names. We categorized complete response and partial response patients as responders. Patients with clinically progressive and stable diseases were marked as non-responders. RNA-seq gene expression profiles of cancer types with clinical response data for fewer than two patients were excluded. At this stage, the filtered data contained gene expression profiles for 29 cancer types. Then for individual cancer types, scaled estimates from gene-level RSEM files were transformed into TPM by multiplying with a factor of 1e6 followed by log2 transformation with the addition of pseudo count 1. For the patients with identical TCGA barcodes, the ones with less non-zero gene expression were selected for downstream analysis.

Drug descriptor data

We obtained drug response information for 192 compounds from the GDSC2 dataset for the 550 cell lines in the CCLE dataset and clinical response information for 215 compounds for 1517 TCGA patients. The chemical structure information for these molecular compounds was retrieved in terms of a simplified molecular-input line-entry system (SMILES) using PubChemPy[24]. However, SMILES were not available for all the molecular compounds. As a result, we ended up with SMILES of 173 and 139 compounds for 550 CCLE cell lines and 1443 unique TCGA patients, respectively. The SMILES2Vec python tool was used to convert these SMILES into vector embeddings by utilizing data of embeddings trained on Pubchem and embedding of size 100[25].

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Pathway activity scores

To train models, we used pathway activity scores. For computing pathway activity scores, we used the Gene Set Variation Analysis (GSVA)[15] R software package to $\log_2(\text{TPM} + 1)$ gene expression matrix and gene set file from Molecular Signatures Database (MSigDB)[48] with min.sz set as 5. We used the c2 collection of canonical pathways (MSigDB.CP.v.6.1) consisting of 1329 gene sets. We integrated the pathway score matrix with the vector embeddings of the drug features. Our final CCLE cell line training dataset constituted 80056 drug/cell line combinations in rows and 1429 features entailing 1329 pathways and drug features of vector size 100 for each molecular compound as the explanatory variable and Z-scores as the response variable in columns. For TCGA patient data, gene expression profiles of individual cancer types were transformed into pathway scores. The GSVA scores of the samples where drug response information was available in each cancer category were merged based on common pathways. The final matrix consisted of 3108 drug/patient combinations and 1427 features (pathways & drug descriptors), and the response variable (responder=1, non-responder=0).

Training models using CCLE RNA-seq cell line dataset

For drug response prediction in gene and pathway space, formulated as a regression task, we used state-of-the-art machine learning methods. We split the CCLE training dataset into an 80% training set and a 20% test set and trained ten bootstrapped regression models for each method using 70% samples as training data. The Random forests were constructed using the ranger R package[49] with 500 trees. ElasticNet was implemented using the glmnet R package[50]. A deep neural network (DNN) was trained using the Keras platform. DNN was modeled with one input layer of as many features detected in the dataset, followed by two hidden layers of sizes 512 and 64, respectively, using RELU as an activation function. Dropout layers were added to avoid the overfitting of models with a dropout rate of 0.2. The ADAM optimizer and Mean squared error as loss function was used. Models were trained using 500 epochs and a batch size of 50. For the model genes as features, the DNN model was trained with an input layer of 600 features and with one hidden layer of 30 using RELU as an activation function. Like for pathway space models,

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ADAM optimizer and Mean squared error as loss function was used in gene space as well. Final predictions were computed as an average of all the predictions from the bootstrapped models.

Training models using TCGA RNA-seq patient profiles

To predict drug response in terms of responder and non-responder, formulated as a classification problem, we utilized the H2O AutoML[36] framework in R for training the TCGA dataset. We split this dataset in the ratio of 80% for training and 20% for testing. This 80% training dataset was subjected to the h2o.automl() function with five-fold cross-validation and a maximum number of models of 20. This resulted in automated training of multiple machine learning models, including Deep learning, DRF, GBM, GLM, XGboost, XRT, and stacked ensemble models.

Survival analysis on TCGA test dataset

The processed TCGA dataset consisted of 3108 drug-patient combinations. We conducted survival analysis on a 20% TCGA test dataset composed of 622 drug-patient entities from 517 unique patients. We stratified the samples using the median value of the predicted probability of the response and computed survival along with 5 year survival probability.

Performance Metrics

We used two accuracy metrics to measure the performance of our models for the regression task: the coefficient of determination (R^2) and Pearson correlation (ρ). R^2 was calculated using the R caret[51] package. For the TCGA dataset, we used MSE, RMSE, LogLoss, Mean Per-Class Error, AUC, AUCPR, and Gini index.

Imputation of missing features

We used an impute[52] package from R to impute if any pathway features are missing in the input test dataset using the nearest neighbor-based averaging approach.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Validating CCLE cell line trained models using scRNA-seq data of cell lines

The scRNA-seq cell line pre-QC UMI count dataset comprising 30314 genes and 56982 cells encompassing 207 cell lines was obtained from Kinker, G. S. et al study[30]. This expression matrix was processed for quality control using an R script (data_processing.R) from the same study. At this stage, we were left with 53299 cells spanning 198 cell lines. This matrix was transformed into TPM by scaling with a factor 1e6 and dividing the UMI count of genes by the total UMI count of the sample. The UMI counts are independent of gene length bias[53]. We log2-transformed this normalized matrix with the addition of a pseudo count of 1. The gene expression values of the same cell lines were averaged. This matrix was converted into pathway scores using GSVA with default parameter settings. Our final validation set consisted of 17279 drug cell line combinations entailing 116 cell lines that overlapped with the cell lines featured in the GDSC2 dataset and tested against 173 drugs.

Predicting the response of paclitaxel drug in MDA-MB-231 single cells using CCLE cell line trained models

Lee et al. dataset was obtained by processing FASTQ files of MDA-MB-231 cells from SRA id SRP040309. We align FASTQ files using STAR aligner[44] with GRCh37 reference genome and GRCh37 GTF file from Ensembl (release 75)[54]. To estimate gene expression, we used HTSeq-count[55]. The HTSeq generated read counts were transformed into TPM values by dividing raw counts by gene length, then scaling with a factor of 1e6 and dividing by total read counts in the sample. We discarded the genes for which gene length was not retrieved by the EDASeq R package[56]. This was followed by the conversion of Ensembl gene IDs to official gene symbols using Homo_sapiens.GRCh37.75.gtf annotation file. We retained genes having a TPM of at least 1 in at least 10% of samples. This matrix was subjected to log2 transformation with the addition of pseudo count 1. We averaged gene expression values of five samples each of naive, stressed, and cells more sensitive to paclitaxel respectively and converted them into pathway scores using GSVA. We predicted drug response for treatment-naive and population of cells more sensitive to paclitaxel in terms of Z-score of LN IC50 values.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Validating TCGA data trained models using BRAF-mutant melanoma patient profiles

We used publicly available Reads Per Kilobase of transcript per million mapped reads (RPKM) RNA-seq profiles of BRAF mutant melanoma patients for drug response prediction. This dataset comprises six paired patients for which RNA-seq was performed before treatment and after treatment with RAF or RAF+MEK inhibitors during disease progression. This dataset includes therapy response information as well. This dataset is referred to as Wagle, Nikhil, et al. dataset[38]. For our analysis, we transformed RPKM normalized RNA-seq profiles to TPM by dividing each RPKM value by the sum of the RPKM values for all genes in the sample and multiplying by one million[57]. The TPM normalized matrix was log2 transformed with the addition of a pseudo count of 1. The resulting matrix was transformed into pathway scores using GSVA. For the three patients for whom information on the mechanism of acquired resistance to dabrafenib/trametinib is available, we predicted a therapeutic response to dabrafenib and trametinib.

Code availability

https://github.com/SmritiChawla/PrecOnco

Author contributions

DS conceived the study with CCN. CCN and BGH contributed to the experimental design of prostate cancer study. SC conducted data analyses with assistance from SP, PR and NB. AJ, AA, AG, ND trained machine learning models. DS, AM and Jayadeva supervised the development of machine learning models. AR supervised and performed significant parts of data analysis and data interpretation. ML contributed to the RNA-seq pipeline of prostate cancer datasets, data interpretation and analysis. ER contributed in experimental design and animal work for LNCaP xenografts study, RNA extraction and sample QC. GA assisted in biological interpretation of results. DS, CCN, BGH supervised the entire study. All authors substantially contributed for manuscript writing and reviewing.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Conflict of interest

None declared.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

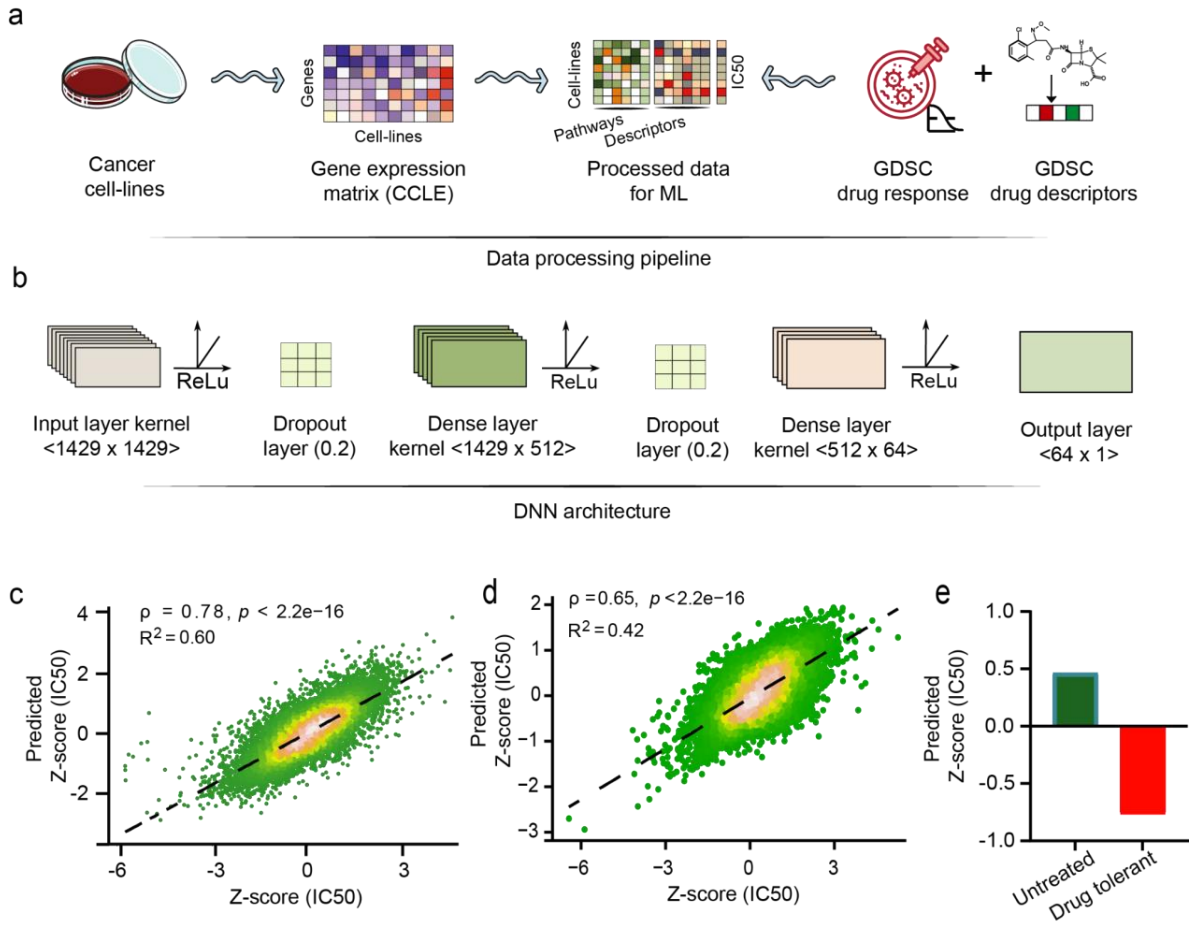


Fig. 1. Illustration of the predictive analysis workflow. a Schematic workflow depicting data processing pipeline. The first step involves the processing of training data. The RNA-seq RSEM TPM genes expression profiles from Cancer Cell Line Encyclopedia (CCLE) were subjected to pathway score transformation using GSVA. This GSVA score matrix was integrated with the drug descriptors obtained in the form of SMILES embedding for each compound. **b** Model architecture. The second step was training of ML model on this data, comprising GSVA scores and drug descriptors as an explanatory variable set and Z-scores of the natural log-transformed IC50 values sourced from the GDSC database as the response variable. A deep neural network (DNN) from the Keras platform was used to perform the regression task of predicting drug response. **c** Scatter plot demonstrating model efficiency on the CCLE bulk RNA-seq test dataset in terms of observed vs predicted Z-score (LN IC50) measured by Pearson correlation (ρ) = 0.78 and coefficient of determination (R^2) = 0.60. **d** Scatter plot demonstrating model efficiency on the Kinker, G. S. et al. scRNA-seq cell lines dataset comparing observed vs predicted Z-score (LN IC50) measured by Pearson correlation (ρ) = 0.65 and coefficient of determination (R^2) = 0.42. **e** Evaluation of predicted drug response of paclitaxel in scRNA-seq MDA-MB-231 cells. Barplots depicting predicted response for paclitaxel in treatment-naive and paclitaxel sensitive cell population.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

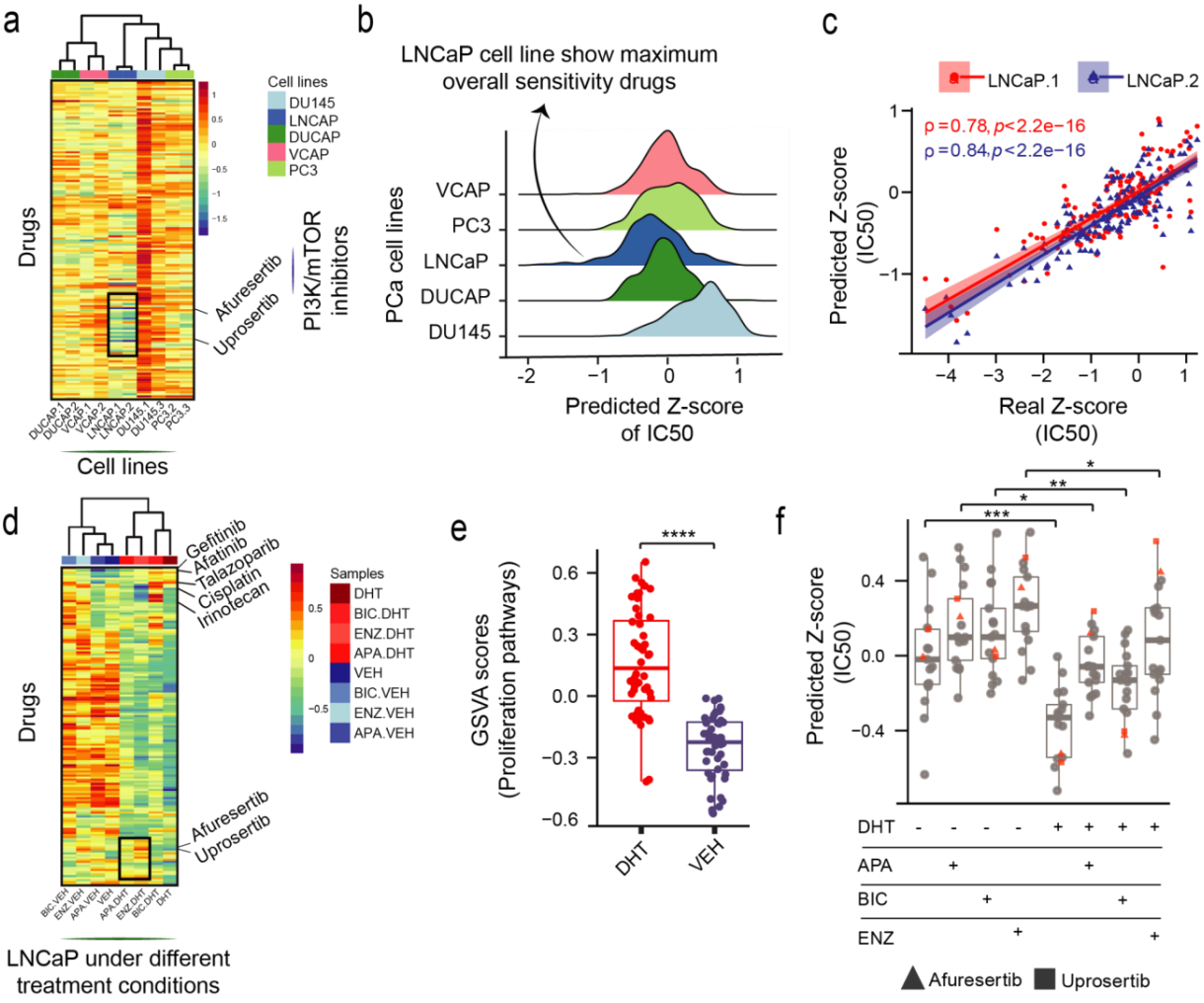
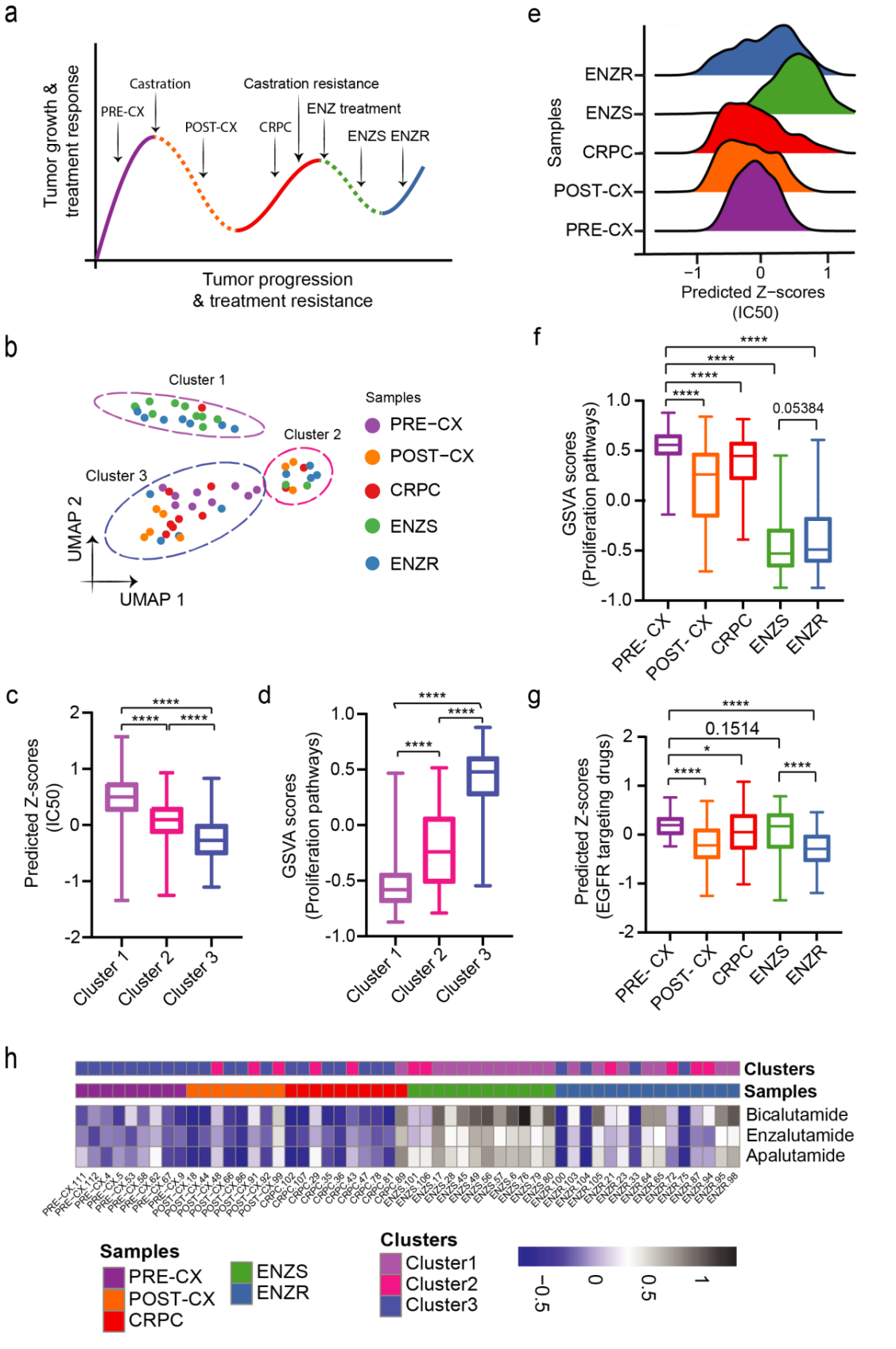


Fig. 2. Analysis of drug response prediction in prostate cancer cell lines a Heatmap showing predicted Z-scores (LN IC50) for 155 drugs across five prostate cancer baseline cell lines highlighting PI3K/mTOR signaling targeting drugs. The lower the Z-score, the more sensitive a sample is predicted to a drug. Color bars indicate cell lines type. **b** Ridgeplot showing overall distribution of predicted Z-scores (LN IC50) across five prostate cancer cell lines. **c** Scatterplot depicting Pearson correlation (*ρ*) between real and predicted Z-scores (IC50) for the LNCaP cell line. The line color indicated two biological replicates of the LNCaP cell line. **d** Heatmap of predicted Z-scores (LN IC50) of LNCaP cells in the presence and absence of androgens (DHT) and AR antagonists (ENZ, BIC, and APA) to 155 drugs. **e** Boxplots depicting the distribution of GSVA scores of proliferation-related pathways in the presence vs. absence of DHT (Asterisks denote Wilcoxon rank sum test *p-value*). **f** Boxplot of predicted Z-scores for drugs targeting PI3K/mTOR signaling across the treatment conditions (Asterisks denotes Wilcoxon rank sum test *p-value*).

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig. 3. Analysis of drug response prediction in LNCaP derived xenografts. a. Overall schematics of the experimental setup of LNCaP xenograft-based prostate cancer progression study with indicated treatments, therapeutic response, and therapeutic resistance stages. Solid lines represent growth and treatment resistance; dotted lines are treatment responsiveness **b** UMAP based projections of predicted Z-scores showing three separate clusters. **c** Box plots depicting distribution predicted Z-score (LN IC50) across 3 clusters (Asterisks denote Wilcoxon rank sum test *p-value*). **d** Boxplots showing the distribution of GSVA scores of proliferation-related pathways across 3 clusters (Asterisks denote Wilcoxon rank sum test *p-value*). **e** Ridgeplot showing overall distribution of predicted Z-scores (LN IC50) across tumor types. **f** Boxplots showing the distribution of GSVA scores of proliferation-related pathways across tumor types (Asterisks denote Wilcoxon rank sum test *p-value*). **g** Box plot depicting predicted Z-scores (LN IC50) of EGFR signaling pathway targeting drugs (Asterisks denote Wilcoxon rank sum test *p-value*). **h** Heatmap showing predicted Z-scores (LN IC50) for three 'unseen' drugs not present in GDSC, namely—BIC, APA, and ENZ. Color bars indicate tumor types and clusters as acquired through UMAP projections.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

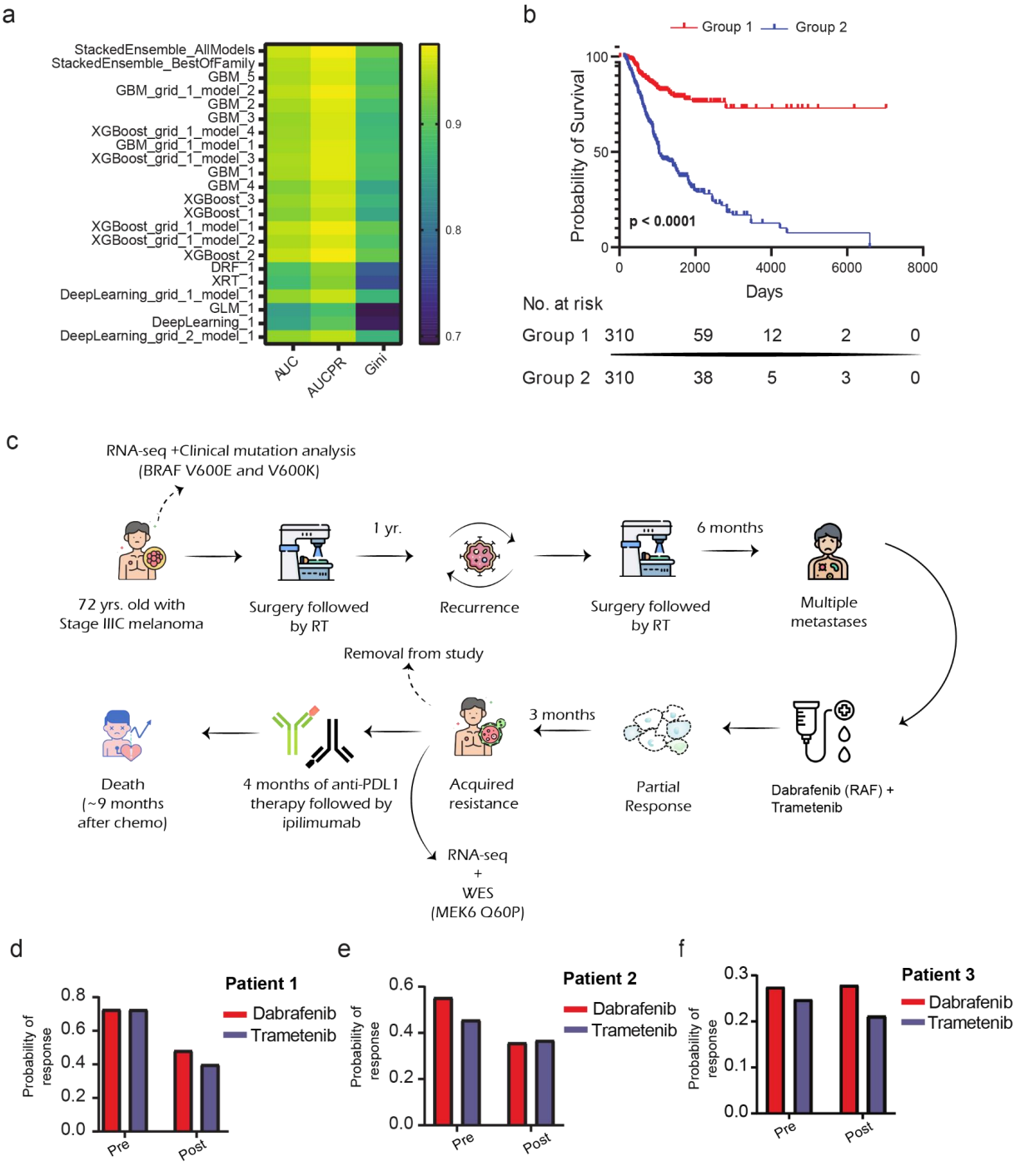


Fig. 4. Evaluation of TCGA model efficiency a Heatmap showing performance metrics, namely AUC, AUCPR, and Gini index for 22 AutoML trained models on the test dataset. **b** Survival analysis of a test dataset spanning multiple TCGA cancer types. Patients were classified into two groups based on the median value of the predicted probability of response with a P-value < 0.0001 (log-rank test). **c** Schematic representation of the journey of melanoma patient 1 from diagnosis of

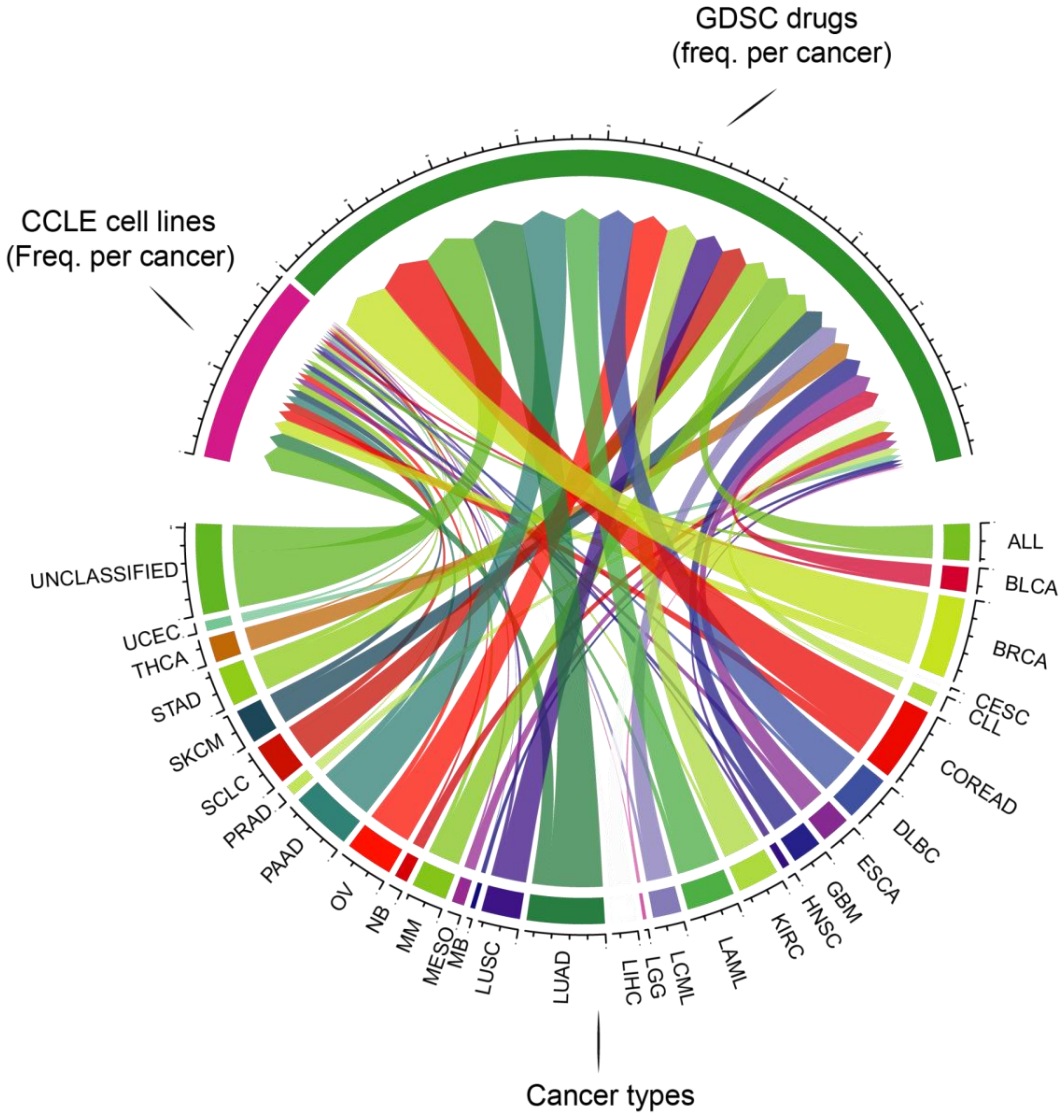
bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

melanoma to initial treatments and enrolment into phase I/II study of dabrafenib and trametinib which resulted in initial regression of disease but ultimately the patient died after discontinuation of chemotherapy. **d-f** Bar plots showing the probability of predicted response for chemotherapy drugs dabrafenib and trametinib in patient 1, patient 2, and patient 3, respectively.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary data

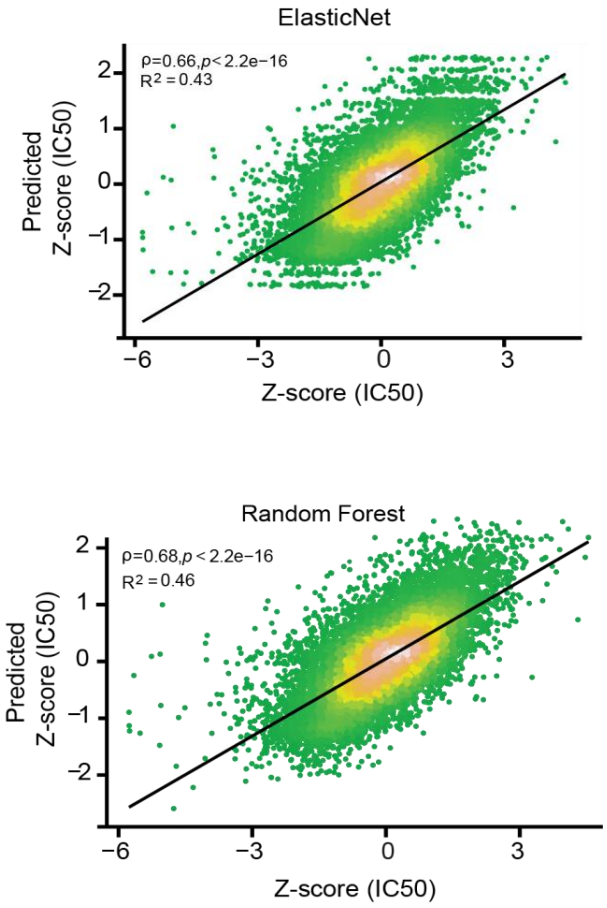
a



Supplementary Fig. 1. Overall summary of the CCLE and GDSC dataset used for training the Deep Neural Network (DNN). a Chord diagram showing the frequency of cell lines (n=550) from the CCLE database along with the frequency of tested drugs (n=173) from the GDSC database spanning 29 classified cancer types used for the training dataset.

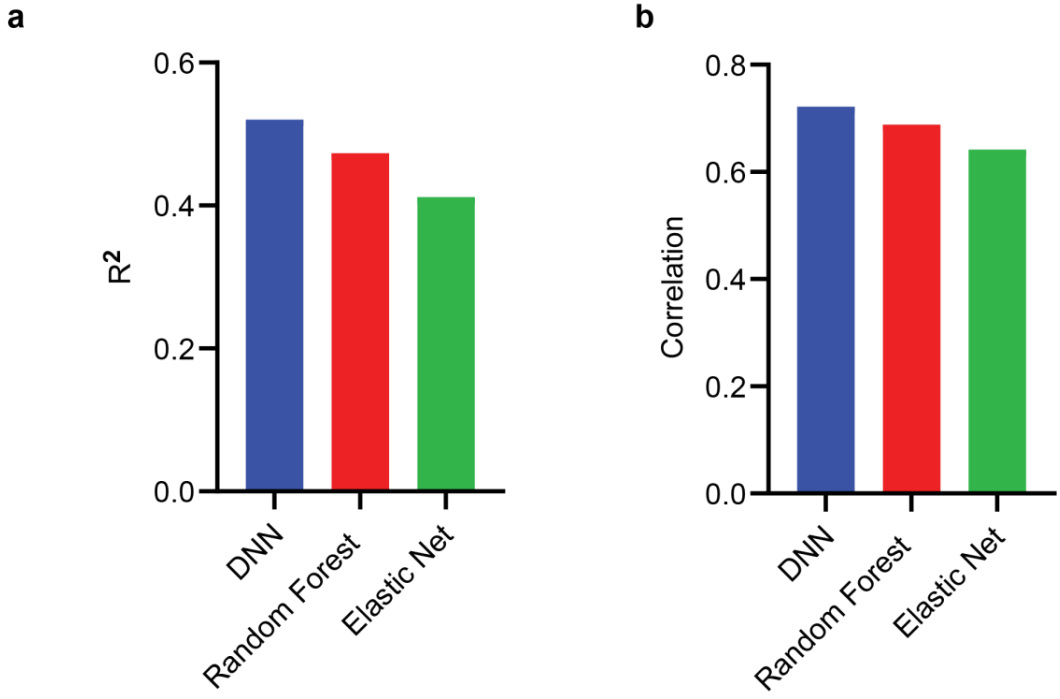
bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

a



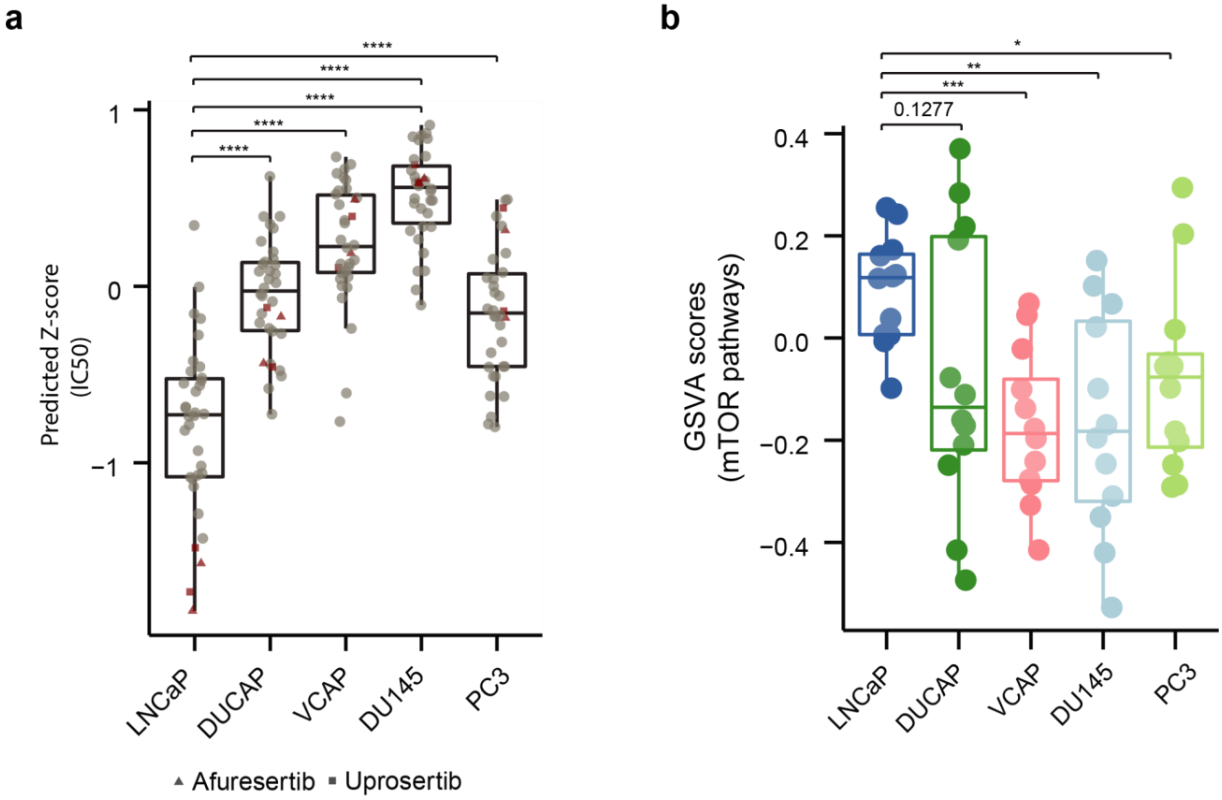
Supplementary Fig. 2. Performance evaluation of different machine learning models trained using CCLE bulk RNA-seq cell lines in pathway space. a Scatter plot of observed and predicted Z-score (IC50) showing prediction performance of Elastic net and Random Forest. We calculated Pearson correlation (ρ) and coefficient of determination (R^2) for evaluation.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



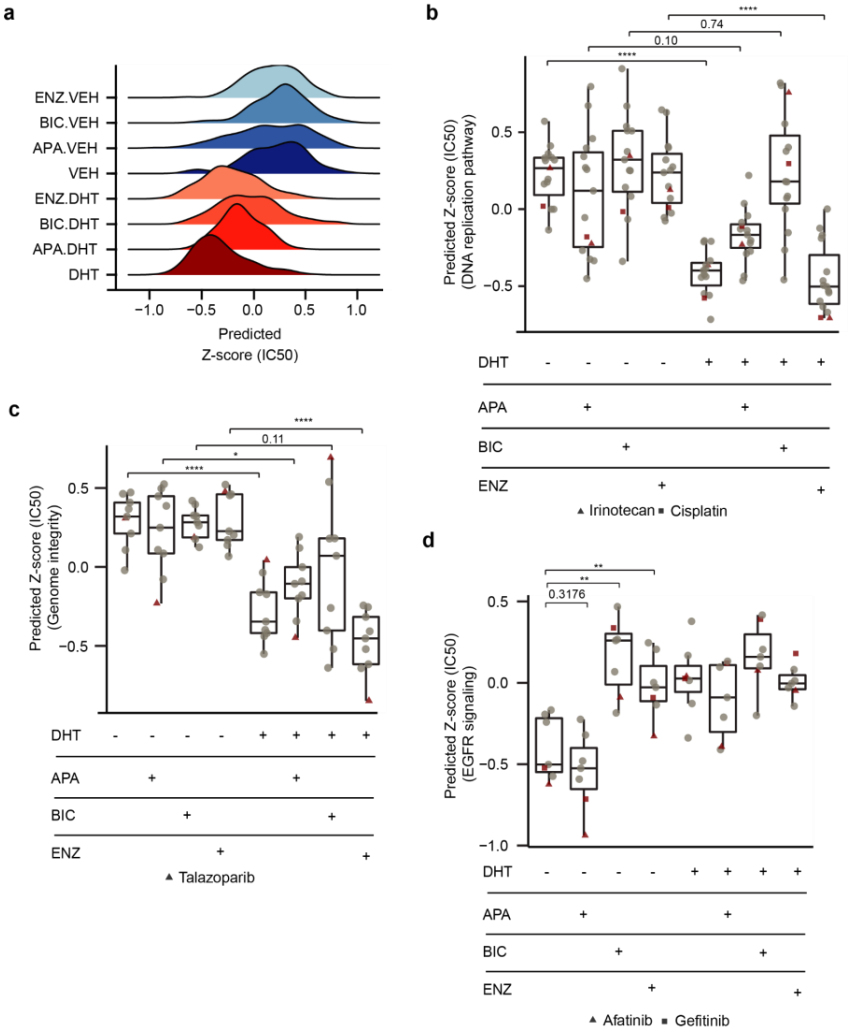
Supplementary Fig. 3. Performance evaluation of different machine learning models trained using CCLE bulk RNA-seq cell lines gene space. a Barplots showing coefficient of determination (R^2). **b** Barplots showing Pearson correlation (ρ).

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



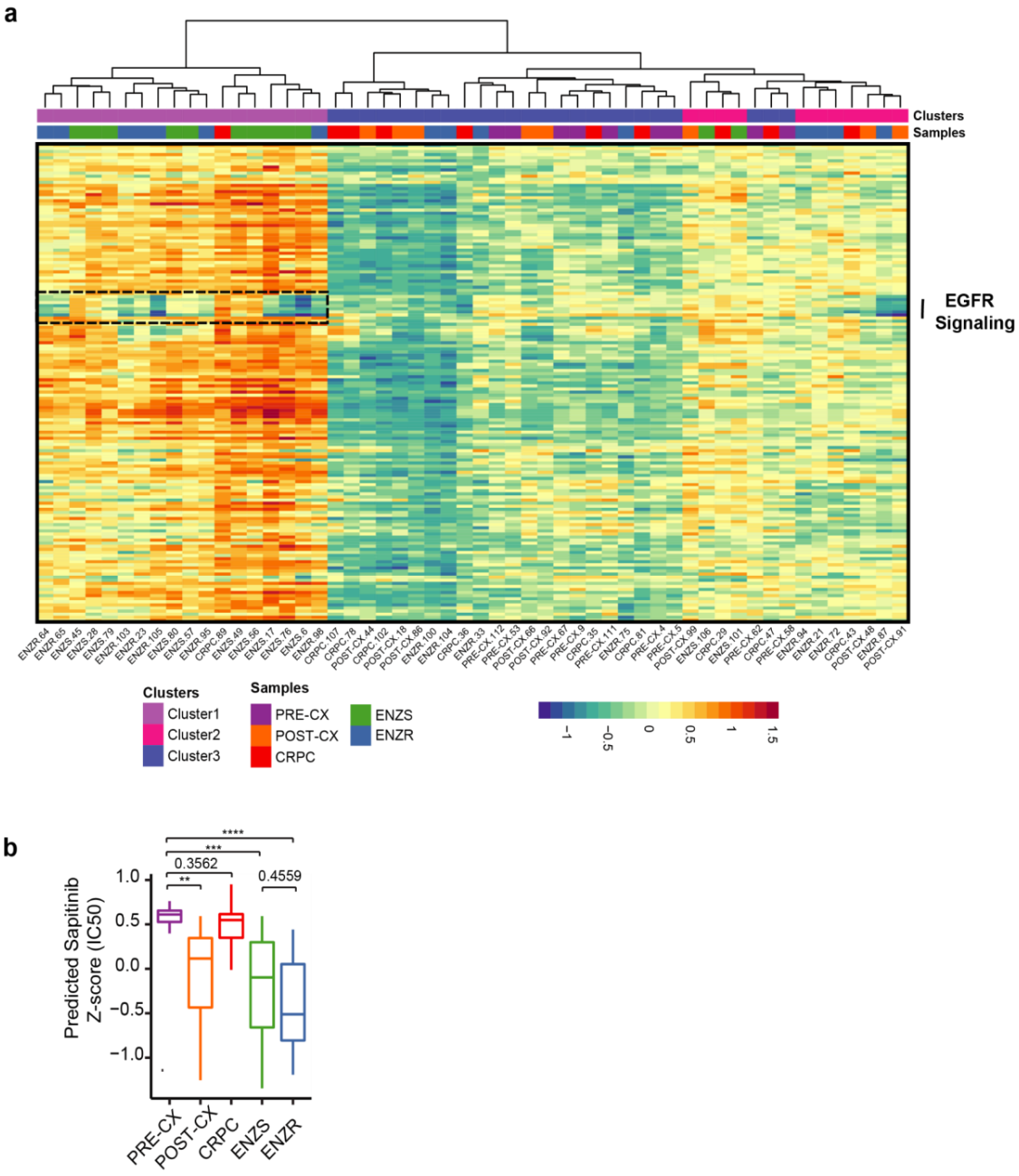
Supplementary Fig. 4. LNCaP cell line sensitivity to drugs targeting mTOR signaling. A Box plots depicting the distribution of drug response prediction of mTOR/PI3K signaling targeting drugs. LNCaP cell line, in particular, was predicted to be more sensitive to these drugs, with afuresertib and uprosertib having the most profound effect (Asterisks denote Wilcoxon rank sum test *p-value*). **b** Box plots showing the distribution of GSVA scores of mTOR-related pathways highlighting increased expression of these pathways in the LNCaP cell line (Asterisks denote Wilcoxon rank sum test *p-value*).

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Supplementary Fig. 5. Drug response prediction and analysis in LNCAP cells under different treatment conditions. a Ridge plot showing the overall distribution of predicted Z-score (IC50) of 155 drugs tested against prostate cancer cell lines in the GDSC2 dataset across the treatment conditions. **b** Boxplot depicting predicted Z-score (IC50) of drugs targeting DNA replication pathway explicitly highlighting irinotecan and cisplatin (Asterisks denote Wilcoxon rank sum test *p-value*). **c** Boxplot showing predicted Z-score (IC50) of drugs targeting genome integrity pathway specifically pinpointing talazoparib (Asterisks denote Wilcoxon rank sum test *p-value*). **d** Boxplot of predicted Z-score (IC50) of the drugs targeting EGFR signaling across the treatment conditions specifically highlighting afatinib and gefitinib drugs (Asterisks denote Wilcoxon rank sum test *p-value*).

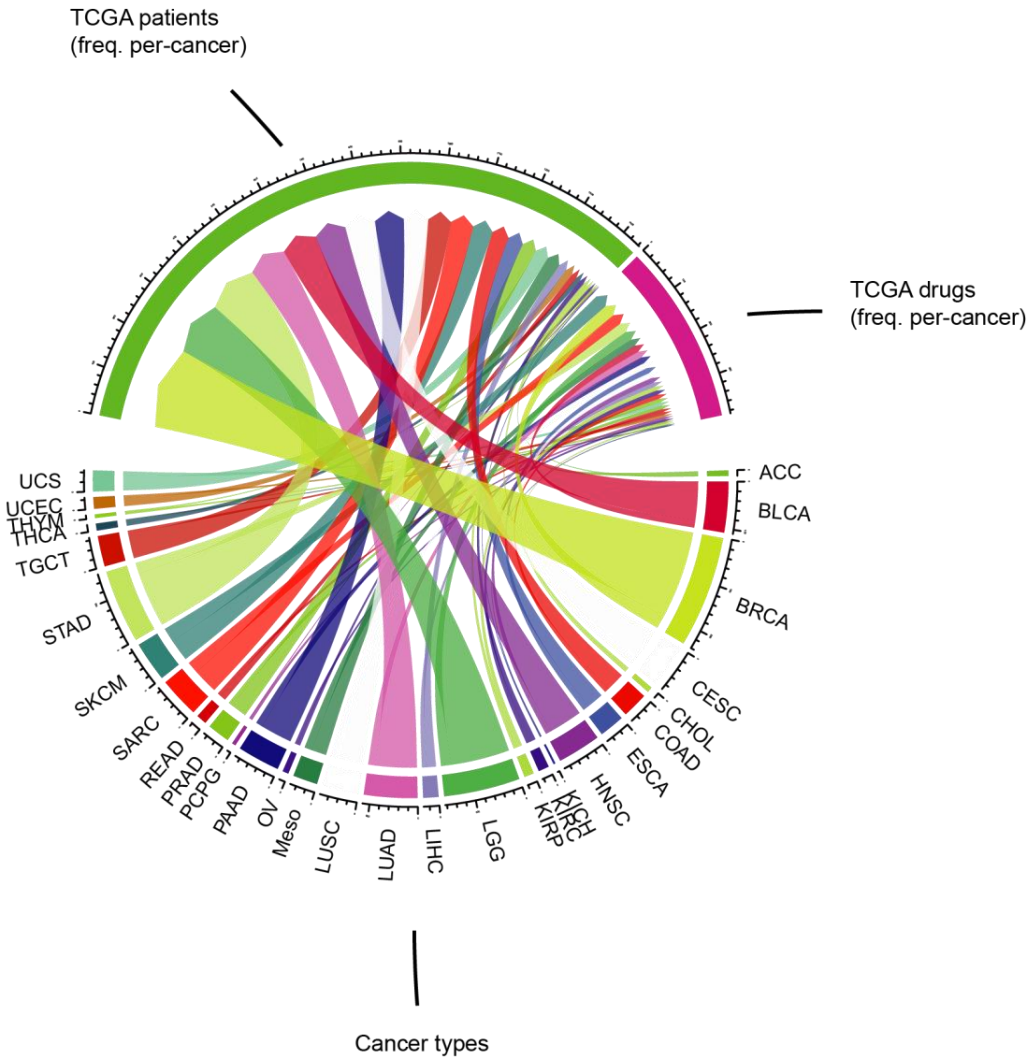
bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Supplementary Fig. 6. Predictions in LNCAP xenografts under different treatment conditions. a Visualization of predicted Z-scores (IC50) in 54 xenograft tumor samples for 155 GDSC drugs. Heatmap showing the sensitivity of some ENZ treated tumors to 'EGFR signaling' targeting drugs. **b** Boxplot of predicted Z-scores (IC50) of sapitinib across the tumor types revealing the highest sensitivity of this drug for ENZR tumors (Asterisks denote Wilcoxon rank sum test *p-value*).

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

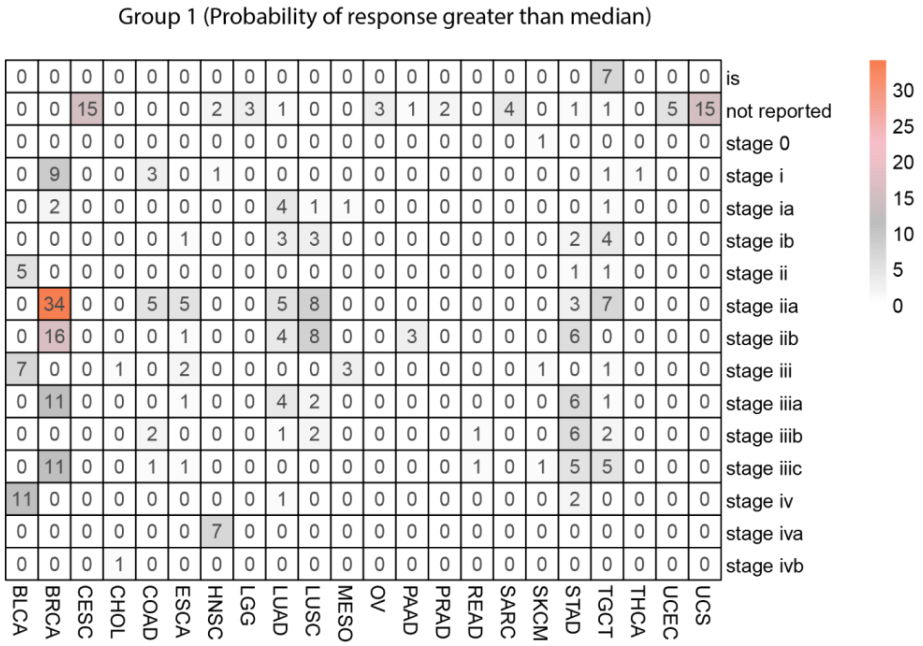
a



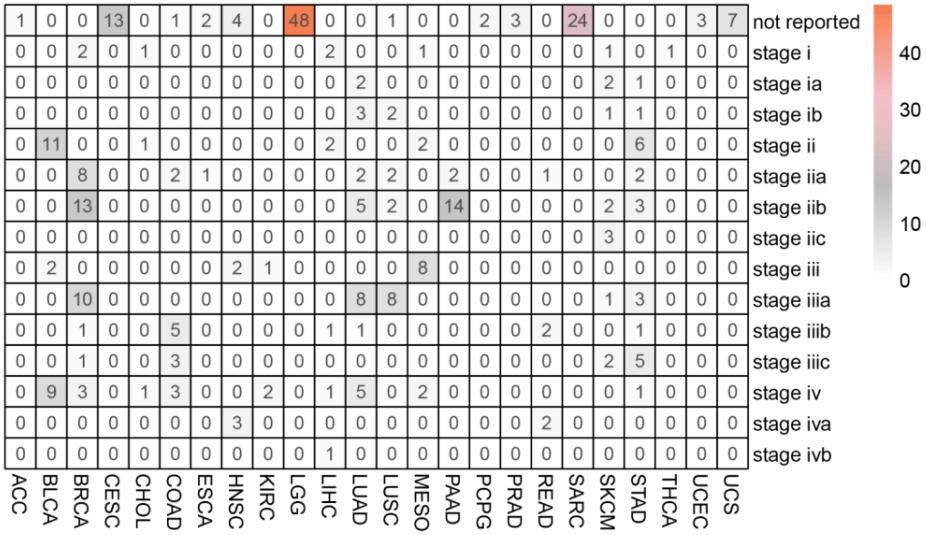
Supplementary Fig. 7. Description of the dataset used for training AutoML models. a Chord diagram showing the frequency of patients (n=1443) from the GDAC firehose database along with the frequency of tested drugs (n=139) from TCGA clinical response data spanning 29 classified cancer types used for the training dataset.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

a

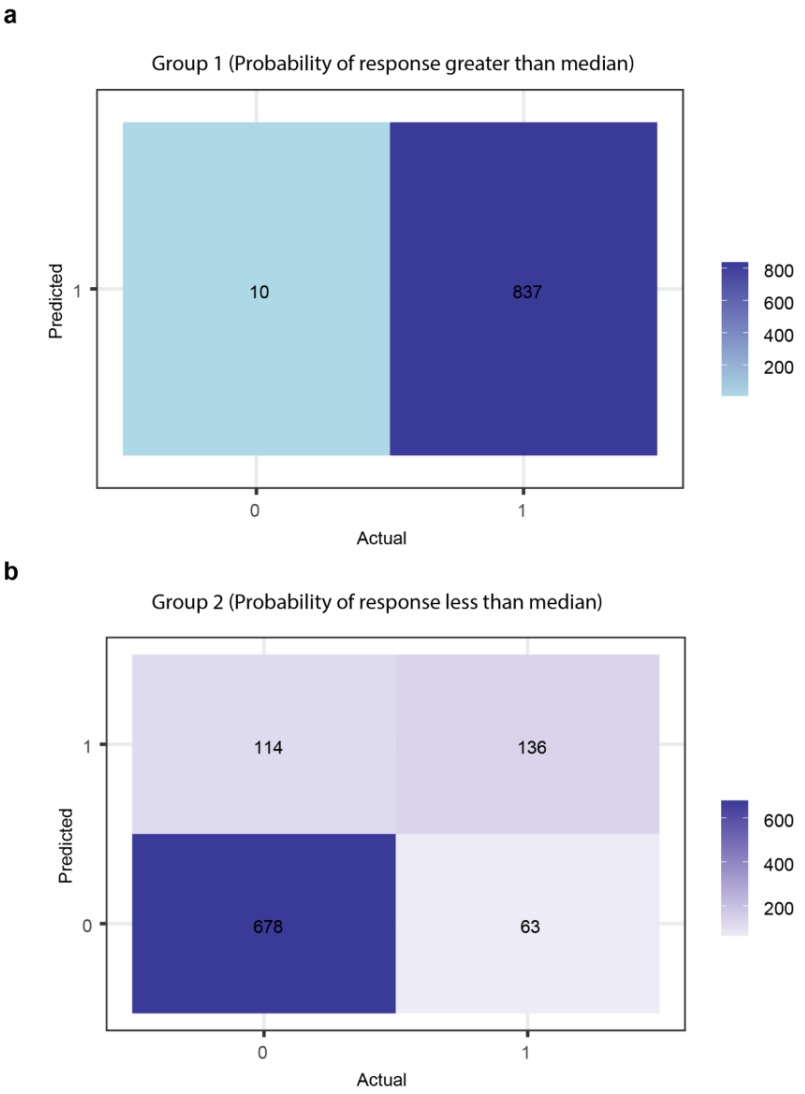


b



Supplementary Fig. 8. Visualization of frequency of stages across different cancer types for TCGA test dataset. a & b Heatmaps showing the frequency distribution of stages across cancer for group 1 and group 2, respectively, stratified based on the median value of the probability of response as predicted through classifier trained on TCGA dataset. We denoted the frequency of stage with 0 if stage is unknown.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Supplementary Fig. 9. Visualization of confusion matrices for classification results of TCGA test dataset. a Heatmap for confusion matrix showing classification results for group 1. We stratified patients into group 1 and group 2 based on the median value of the probability of response as predicted through a classifier trained on TCGA dataset. Our median value of 0.80 yielded only patients belonging to class 1 (responder) in group 1. **b** Heatmap for confusion matrix showing classification results for group 2. The actual class depicts treatment response information from TCGA clinical metadata files. The treatment response information is categorized as responder (label=1) for complete response and partial response of patients. While clinical progressive disease and stable disease in patients represent non-responder categories (label=0).

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Table 1. Performance summary of AutoML models on the bulk RNA-seq TCGA patient test dataset

Models	MSE	RMSE	Log Loss	Mean per class error	AUC	AUCPR	Gini
StackedEnsemble_AllModels_AutoML_20210812_215305	0.08	0.28	0.27	0.1	0.96	0.98	0.91
StackedEnsemble_BestOfFamily_AutoML_20210812_215305	0.09	0.29	0.29	0.1	0.95	0.97	0.9
GBM_5_AutoML_20210812_215305	0.09	0.3	0.31	0.11	0.94	0.96	0.89
GBM_grid__1_AutoML_20210812_215305_model_2	0.09	0.3	0.29	0.1	0.95	0.97	0.9
GBM_2_AutoML_20210812_215305	0.1	0.31	0.31	0.11	0.94	0.96	0.89
GBM_3_AutoML_20210812_215305	0.1	0.32	0.33	0.15	0.94	0.96	0.87
XGBoost_grid__1_AutoML_20210812_215305_model_4	0.1	0.31	0.31	0.12	0.94	0.96	0.88
GBM_grid__1_AutoML_20210812_215305_model_1	0.1	0.32	0.33	0.13	0.94	0.97	0.88
XGBoost_grid__1_AutoML_20210812_215305_model_3	0.09	0.3	0.3	0.11	0.94	0.97	0.89
GBM_1_AutoML_20210812_215305	0.1	0.31	0.32	0.13	0.94	0.97	0.89
GBM_4_AutoML_20210812_215305	0.11	0.33	0.34	0.14	0.93	0.96	0.85
XGBoost_3_AutoML_20210812_215305	0.1	0.32	0.32	0.15	0.93	0.96	0.87
XGBoost_1_AutoML_20210812_215305	0.11	0.33	0.34	0.17	0.93	0.95	0.85
XGBoost_grid__1_AutoML_20210812_215305_model_1	0.08	0.29	0.28	0.13	0.95	0.97	0.91
XGBoost_grid__1_AutoML_20210812_215305_model_2	0.09	0.3	0.3	0.11	0.95	0.96	0.89
XGBoost_2_AutoML_20210812_215305	0.08	0.29	0.28	0.1	0.95	0.97	0.91
DRF_1_AutoML_20210812_215305	0.14	0.38	0.44	0.22	0.89	0.93	0.78
XRT_1_AutoML_20210812_215305	0.15	0.38	0.45	0.23	0.88	0.93	0.76
DeepLearning_grid__1_AutoML_20210812_215305_model_1	0.11	0.33	0.52	0.12	0.93	0.96	0.87
GLM_1_AutoML_20210812_215305	0.15	0.39	0.47	0.24	0.85	0.89	0.69
DeepLearning_1_AutoML_20210812_215305	0.16	0.41	0.53	0.23	0.85	0.9	0.7
DeepLearning_grid__2_AutoML_20210812_215305_model_1	0.1	0.31	0.42	0.12	0.93	0.95	0.87

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Table 2. Performance summary of AutoML models on bulk RNA-seq TCGA patient profiles training dataset (cross-validation)

Models	MSE	RMSE	Log Loss	Mean per class error	AUC	AUCPR	Gini
StackedEnsemble_AllModels_AutoML_20210812_215305	0.11	0.34	0.37	0.2	0.91	0.94	0.82
StackedEnsemble_BestOfFamily_AutoML_20210812_215305	0.12	0.34	0.37	0.19	0.91	0.94	0.81
GBM_5_AutoML_20210812_215305	0.12	0.35	0.38	0.17	0.9	0.93	0.81
GBM_grid__1_AutoML_20210812_215305_model_2	0.12	0.35	0.39	0.18	0.9	0.93	0.8
GBM_2_AutoML_20210812_215305	0.13	0.35	0.4	0.19	0.9	0.93	0.8
GBM_3_AutoML_20210812_215305	0.13	0.35	0.4	0.18	0.9	0.93	0.79
XGBoost_grid__1_AutoML_20210812_215305_model_4	0.13	0.36	0.41	0.21	0.89	0.92	0.79
GBM_grid__1_AutoML_20210812_215305_model_1	0.13	0.36	0.4	0.19	0.89	0.93	0.79
XGBoost_grid__1_AutoML_20210812_215305_model_3	0.13	0.35	0.4	0.17	0.89	0.93	0.79
GBM_1_AutoML_20210812_215305	0.13	0.36	0.41	0.19	0.89	0.93	0.79
GBM_4_AutoML_20210812_215305	0.13	0.36	0.42	0.19	0.89	0.93	0.78
XGBoost_3_AutoML_20210812_215305	0.13	0.36	0.42	0.19	0.89	0.92	0.78
XGBoost_1_AutoML_20210812_215305	0.13	0.36	0.43	0.2	0.88	0.91	0.77
XGBoost_grid__1_AutoML_20210812_215305_model_1	0.13	0.37	0.42	0.21	0.88	0.91	0.77
XGBoost_grid__1_AutoML_20210812_215305_model_2	0.13	0.36	0.43	0.22	0.88	0.91	0.77
XGBoost_2_AutoML_20210812_215305	0.13	0.36	0.43	0.21	0.88	0.91	0.76
DRF_1_AutoML_20210812_215305	0.15	0.39	0.47	0.21	0.87	0.9	0.74
XRT_1_AutoML_20210812_215305	0.15	0.39	0.47	0.24	0.87	0.9	0.73
DeepLearning_grid__1_AutoML_20210812_215305_model_1	0.19	0.43	0.7	0.28	0.82	0.87	0.64
GLM_1_AutoML_20210812_215305	0.17	0.41	0.52	0.28	0.81	0.85	0.62
DeepLearning_1_AutoML_20210812_215305	0.19	0.43	0.67	0.31	0.8	0.85	0.6
DeepLearning_grid__2_AutoML_20210812_215305_model_1	0.18	0.43	0.59	0.33	0.79	0.85	0.58

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary note 1: Clinical characteristics of melanoma patients

Patient 2

Patient 2, a 48-year old man initially diagnosed with stage IB melanoma developed extensively metastatic melanoma after five years of initial diagnosis, confirmed by a pleural biopsy (pre-treatment). Further, clinical mutation analysis revealed the presence of BRAF V600E mutation. The patient was subjected to first-line treatment of Dabrafenib and Trametinib which showed partial response but after 3 months, routine scans revealed significant disease progression. The potential cause of resistance to therapy was the presence of BRAF splice variant as detected by RNA-seq and whole exome sequencing (WES) in post-treatment tumors but not in pre-treatment tumors. The patient died six months after being diagnosed with metastatic disease[38].

Patient 3

Patient 3, a 42-year old man underwent surgery and lymph node dissection of stage IIIC melanoma of the left thigh. The patient had a BRAF V600E mutation as revealed by clinical mutational analysis. After six months of surgery, the patient was subjected to first-line therapy of dabrafenib and trametinib. But nearly after one year the patient developed progressive disease and the potential underlying cause of acquired resistance to therapy was the presence of BRAF amplification in post-treatment tumors as depicted by WES. The patient was given Ipilimumab for a short time but died after four cycles and three months after discontinuation of chemotherapy[38].

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1. Gerdes, H. *et al.* Drug ranking using machine learning systematically predicts the efficacy of anti-cancer drugs. *Nat. Commun.* **12**, 1850 (2021).
2. Feng, F., Shen, B., Mou, X., Li, Y. & Li, H. Large-scale pharmacogenomic studies and drug response prediction for personalized cancer medicine. *J. Genet. Genomics* (2021) doi:10.1016/j.jgg.2021.03.007.
3. Baudino, T. A. Targeted Cancer Therapy: The Next Generation of Cancer Treatment. *Curr. Drug Discov. Technol.* **12**, 3–20 (2015).
4. Adam, G. *et al.* Machine learning approaches to drug response prediction: challenges and recent progress. *NPJ Precis Oncol* **4**, 19 (2020).
5. Maeda, H. & Khatami, M. Analyses of repeated failures in cancer therapy for solid tumors: poor tumor-selective drug delivery, low therapeutic efficacy and unsustainable costs. *Clin. Transl. Med.* **7**, 11 (2018).
6. Suphavilai, C., Bertrand, D. & Nagarajan, N. Predicting Cancer Drug Response using a Recommender System. *Bioinformatics* **34**, 3907–3914 (2018).
7. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).
8. Yang, W. *et al.* Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* **41**, D955–61 (2013).

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

9. Jia, P. *et al.* Deep generative neural network for accurate drug response imputation. *Nat. Commun.* **12**, 1740 (2021).

10. Cancer Genome Atlas Research Network *et al.* The Cancer Genome Atlas Pan-Cancer analysis project. *Nat. Genet.* **45**, 1113–1120 (2013).

11. Ammad-Ud-Din, M. *et al.* Drug response prediction by inferring pathway-response associations with kernelized Bayesian matrix factorization. *Bioinformatics* **32**, i455–i463 (2016).

12. Wang, L., Li, X., Zhang, L. & Gao, Q. Improved anticancer drug response prediction in cell lines using matrix factorization with similarity regularization. *BMC Cancer* **17**, 513 (2017).

13. Chang, Y. *et al.* Cancer Drug Response Profile scan (CDRscan): A Deep Learning Model That Predicts Drug Effectiveness from Cancer Genomic Signature. *Sci. Rep.* **8**, 8857 (2018).

14. Sakellaropoulos, T. *et al.* A Deep Learning Framework for Predicting Response to Therapy in Cancer. *Cell Rep.* **29**, 3367–3373.e4 (2019).

15. Hänzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* **14**, 7 (2013).

16. Chawla, S. *et al.* UniPath: a uniform approach for pathway and gene-set based analysis of heterogeneity in single-cell epigenome and transcriptome profiles. *Nucleic Acids Res.* **49**, 1801 (2021).

17. Li, H. *et al.* Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat. Genet.* **49**, 708–718 (2017).

18. Ben-Hamo, R. *et al.* Predicting and affecting response to cancer therapy based on pathway-level biomarkers. *Nat. Commun.* **11**, 3296 (2020).

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

19. Suphavilai, C., Chia, S., Sharma, A., Tu, L. & Peres, R. Predicting heterogeneity in clone-specific therapeutic vulnerabilities using single-cell transcriptomic signatures. *bioRxiv* (2020).

20. Teo, M. Y., Rathkopf, D. E. & Kantoff, P. Treatment of Advanced Prostate Cancer. *Annu. Rev. Med.* **70**, 479–499 (2019).

21. Komura, K. *et al.* Current treatment strategies for advanced prostate cancer. *Int. J. Urol.* **25**, 220–231 (2018).

22. Swami, U., McFarland, T. R., Nussenzveig, R. & Agarwal, N. Advanced Prostate Cancer: Treatment Advances and Future Directions. *Trends Cancer Res.* **6**, 702–715 (2020).

23. Gupta, S. *et al.* Prioritization of anticancer drugs against a cancer using genomic features of cancer cells: A step towards personalized medicine. *Sci. Rep.* **6**, 23857 (2016).

24. Swain, M. PubChemPy: A way to interact with PubChem in Python. (2014).

25. Öztürk, H., Ozkirimli, E. & Özgür, A. A novel methodology on distributed representations of proteins using their interacting ligands. *Bioinformatics* **34**, i295–i303 (2018).

26. Chollet, F. & Others. keras. (2015).

27. Riddick, G. *et al.* Predicting in vitro drug sensitivity using Random Forests. *Bioinformatics* **27**, 220–224 (2011).

28. Jang, I. S., Neto, E. C., Guinney, J., Friend, S. H. & Margolin, A. A. Systematic assessment of analytical methods for drug sensitivity prediction from cancer cell line data. *Pac. Symp. Biocomput.* 63–74 (2014).

29. Ding, M. Q., Chen, L., Cooper, G. F., Young, J. D. & Lu, X. Precision Oncology beyond Targeted Therapy: Combining Omics Data with Machine Learning Matches the Majority of

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Cancer Cells to Effective Therapeutics. *Mol. Cancer Res.* **16**, 269–278 (2018).

30. Kinker, G. S. *et al.* Pan-cancer single-cell RNA-seq identifies recurring programs of cellular heterogeneity. *Nat. Genet.* **52**, 1208–1218 (2020).
31. Lee, M.-C. W. *et al.* Single-cell analyses of transcriptional heterogeneity during drug tolerance transition in cancer cells by RNA sequencing. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E4726–35 (2014).
32. Mateo, J. *et al.* Accelerating precision medicine in metastatic prostate cancer. *Nat Cancer* **1**, 1041–1053 (2020).
33. Nevedomskaya, E., Baumgart, S. J. & Haendler, B. Recent Advances in Prostate Cancer Treatment and Drug Discovery. *Int. J. Mol. Sci.* **19**, (2018).
34. Heinlein, C. A. & Chang, C. Androgen receptor in prostate cancer. *Endocr. Rev.* **25**, 276–308 (2004).
35. Rajamahanty, S., Alonzo, C., Aynehchi, S., Choudhury, M. & Konno, S. Growth inhibition of androgen-responsive prostate cancer cells with brefeldin A targeting cell cycle and androgen receptor. *J. Biomed. Sci.* **17**, 5 (2010).
36. LeDell, E. & Poirier, S. H2o automl: Scalable automatic machine learning. in *Proceedings of the AutoML Workshop at ICML* vol. 2020 (2020).
37. Hawkes, N. Cancer survival data emphasise importance of early diagnosis. *BMJ* **364**, l408 (2019).
38. Wagle, N. *et al.* MAP kinase pathway alterations in BRAF-mutant melanoma patients with acquired resistance to combined RAF/MEK inhibition. *Cancer Discov.* **4**, 61–68 (2014).
39. Eisenhauer, E. A. *et al.* New response evaluation criteria in solid tumours: revised RECIST

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

guideline (version 1.1). *Eur. J. Cancer* **45**, 228–247 (2009).

40. Yu, K., Toral-Barza, L., Shi, C., Zhang, W.-G. & Zask, A. Response and determinants of cancer cell susceptibility to PI3K inhibitors: combined targeting of PI3K and Mek1 as an effective anticancer strategy. *Cancer Biol. Ther.* **7**, 307–315 (2008).
41. Steele, T. M., Jathal, M. K., Siddiqui, S. & Ghosh, P. M. Abstract 4676: Overcoming EGFR-induced resistance to enzalutamide in castration-resistant prostate cancer. *Cancer Res.* **76**, 4676–4676 (2016).
42. Andrews, S. Babraham bioinformatics-FastQC a quality control tool for high throughput sequence data. *URL: https://www. bioinformatics. babraham. ac. uk/projects/fastqc* (2010).
43. Krueger, F. Trim Galore: a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files, with some extra functionality for MspI-digested RRBS-type (Reduced Representation Bisufite-Seq) libraries. *URL http://www. bioinformatics. babraham. ac. uk/projects/trim_galore/. (Date of access: 28/04/2016)* (2012).
44. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
45. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
46. Broad Institute TCGA Genome Data Analysis Center. Analysis-ready standardized TCGA data from Broad GDAC Firehose 2016_01_28 run. (2016) doi:10.7908/C11G0KM9.
47. Grossman, R. L. *et al.* Toward a Shared Vision for Cancer Genomic Data. *N. Engl. J. Med.* **375**, 1109–1112 (2016).

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.17.468905; this version posted November 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

48. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).

49. Wright, M. N. & Ziegler, A. ranger: A Fast Implementation of Random Forests for High Dimensional Data in C and R. *Journal of Statistical Software* vol. 77 (2017).

50. Friedman, J., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* **33**, 1–22 (2010).

51. Kuhn, M. caret: Classification and Regression Training. *Astrophysics Source Code Library* ascl:1505.003 (2015).

52. Hastie, T., Tibshirani, R., Narasimhan, B. & Chu, G. impute: impute: Imputation for microarray data. R package version 1.62. 0. (2020).

53. Phipson, B., Zappia, L. & Oshlack, A. Gene length and detection bias in single cell RNA sequencing protocols. *F1000Res.* **6**, 595 (2017).

54. Howe, K. L. *et al.* Ensembl 2021. *Nucleic Acids Res.* **49**, D884–D891 (2021).

55. Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).

56. Risso, D., Schwartz, K., Sherlock, G. & Dudoit, S. GC-content normalization for RNA-Seq data. *BMC Bioinformatics* **12**, 480 (2011).

57. Zhao, S., Ye, Z. & Stanton, R. Misuse of RPKM or TPM normalization when comparing across samples and sequencing protocols. *RNA* **26**, 903–909 (2020).

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/356908031

Grey Wolf Optimizer-Based Array Reconfiguration to Enhance Power Production from Solar Photovoltaic Plants under Different Scenarios

Article *in* Sustainability · December 2021

DOI: 10.3390/su132413627

CITATIONS
0

READS
20

4 authors, including:

Astitva Kumar
Delhi Technological University
13 PUBLICATIONS **30** CITATIONS

SEE PROFILE

Mohammad Rizwan
Delhi Technological University
13 PUBLICATIONS **36** CITATIONS

SEE PROFILE

Uma Nangia
Delhi Technological University
33 PUBLICATIONS **203** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Smart Grid Optimization View project

Optimal Design of SPV systems View project

All content following this page was uploaded by Astitva Kumar on 09 December 2021.

The user has requested enhancement of the downloaded file.

sustainability

MDPI

Article

Grey Wolf Optimizer-Based Array Reconfiguration to Enhance Power Production from Solar Photovoltaic Plants under Different Scenarios

Astitva Kumar [1,*], Mohammad Rizwan [1], Uma Nangia [1] and Muhannad Alaraj [2]

[1] Department of Electrical Engineering, Delhi Technological University, Delhi 110042, India; rizwan@dce.ac.in (M.R.); umanangia@dce.ac.in (U.N.)
[2] Department of Electrical Engineering, College of Engineering, Qassim University, Buraydah 52571, Saudi Arabia; Muhannad@qu.edu.sa
* Correspondence: astitvakumar_phd2k16@dtu.ac.in

Abstract: The extraction of maximum power is a big challenge in solar photovoltaic-based power plants due to varying atmospheric and meteorological parameters. The concept of array reconfiguration is applied for the maximum power extraction in solar PV plants. Using this approach, the occurrence of multiple peaks in P-V and I-V characteristics during partial shade can be smoothened and reduced significantly. Partial shading due to the movement of the cloud is considered in the research. The cloud movement mainly because of velocity and wind direction is used for creating various shading conditions. The main focus is to reduce the power losses during partial shading using a nature-inspired optimization approach to reconfigure the array for different types of shading conditions. A grey wolf optimizer-based bridge-linked total cross-tied (GWO-BLTCT) configuration is proposed in this paper. The performance of the proposed topology is compared with standard and hybrid topologies, namely, series-parallel, total cross-tied, BLTCT, and SuDoKu-BLTCT, based on performance indicators such as fill factor, performance ratio, power enhancement, and power loss. The proposed GWO-BLTCT outperforms the remaining topologies due to the least power loss and high fill factor. It also has the highest average power enhancement and performance ratio with 23.75% and 70.02% respectively.

Keywords: array reconfiguration; maximum power point; optimization algorithm; partial shading conditions; solar PV system

Citation: Kumar, A.; Rizwan, M.; Nangia, U.; Alaraj, M. Grey Wolf Optimizer-Based Array Reconfiguration to Enhance Power Production from Solar Photovoltaic Plants under Different Scenarios. *Sustainability* **2021**, *13*, 13627. https://doi.org/10.3390/su132413627

Academic Editor: Mohammad Jafari

Received: 18 November 2021
Accepted: 7 December 2021
Published: 9 December 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

In this era, comfort level and convenience are the basic requirements of human beings and should not be compromised at any level. It is expected that the existing and new buildings may have such provision by incorporating the energy efficiency concept or providing the energy-efficient material in the newly constructed buildings. The world's 40% basic resources are being utilized at a rapid pace by the worldwide construction sector. The European Commission initiated the Asia-Link program to impart knowledge about sustainability and its correlation with energy. Environmental sustainability, energy security, and equity create a three-dimensional aspect for energy sustainability, this three-dimensional aspect is known as a trilemma. European countries maintain the top rank in the energy sustainability aspect, whereas India ranks 88th globally [1]. The primary goal of sustainable design is the reduction in the depletion of critical resources, minimizing environmental degradation, and promoting a built environment that is safe, efficient, and productive. These sustainable developments are achieved by promoting the use of renewable energy and developing infrastructure for the vertical expansion of the urban landscape.

Renewable energy (RE) presently has an over 26% share of worldwide electricity production. Solar photovoltaic (SPV) has been the main focus of research and application among renewable energy sources. SPV technologies have boomed mainly due to the cheaper cost of materials, easy availability of solar energy, evolving methods of application of models in residential/commercial sectors, and the involvement of the government and their initiatives. A report was presented by S. Ong et al. for the solar land usage metrics in the USA considering SPV and concentrating solar power (CSP) facilities. Some of the drawbacks of the SPV system are listed below, the land used for the PV system compared with CSP is 8.4% more [2]. Furthermore, the PV arrays have a low conversion efficiency of with 24% being the highest for the most common silicon crystalline module [3]. It becomes even more critical and challenging to improve its performance in different conditions.

Partial shading (PS) is a condition when solar panels are exposed to uneven solar irradiance. PS is temporary and nonlinear. The partial shading condition (PSC) results in a lower maximum power point (MPP) and is commonly referred to as a mismatch fault. This condition arises due to various factors like moving clouds, nearby buildings, poles, trees, and other factors [4]. The adverse impact of the PSCs can be listed as (a) hot spot formation leads to permanent damage of the PV panels; (b) multiple peaks of the MPP that create excess stress on the tracking of the MPP. To effectively solve the problem of multiple peaks of power in the P-V curve, the array reconfiguration technique is anticipated. The methodology of PV array reconfiguration is gaining importance lately as it can address the issue of MPP as well as the problem of solar land usage when incorporated in the vertical expansion of the urban landscape [5,6].

Reconfiguration of the PV arrays aims to redistribute the shade to create a uniform row current or equal irradiation. The reconfiguration can be performed either via physical relocation (PR) [7] or via electrical routing of the arrays i.e., electrical array reconfiguration (EAR) [8–10]. PR methods suffer from excessive interconnecting ties and skilled labor [11]. EAR techniques are also the dynamic reconfiguration technique requiring sensors to determine the partial shading and faulty conditions, whereas, in PR methods fixed interconnection of the PV modules is performed according to the physical location. The significant difference among them is the use of sensors and thus the cost of the system. Several techniques have been proposed to work effectively in tracking the GMPP of PV arrays with PS [12–16]. Different configurations of PV array reconfiguration such as series-parallel (SP), series (S), parallel (P), total cross-tied (TCT), bridge link (BL), and honeycomb (HC) are analyzed in [17] and performance has also been compared. The hybrid configuration of the PV system is superior when compared to the conventional topologies based on power maximization [18]. The topologies derived from SuDoKu game theory have limitations that it can be implemented where the PV array has an even number of rows [19]. In a real-world scenario, the shadow on the PV panels is rarely in the straight lines or a linear pattern, the cloud movement and its modeling are presented by S. Vijayalekshmy et al. for a TCT and rearranged TCT topology [20]. However, the predominant factor for cloud movement and cloud shade due to wind was not taken into consideration. G Sagar et al. had modeled the movement of clouds using wind velocity and had further implemented SuDoKu based shade dispersion through PR of 6 × 6 panels in a BLTCT configuration [21]. Different techniques have been proposed in the literature for shade dispersion through reconfiguration such as magic square (MS), latin square (LS), and dominance square (DS). R. Venkateswari et al. had proposed Lo Shu arrangement for physical reconfiguration of PV array for different types of shading conditions [22]. It was concluded that these were better than conventional topologies. Tables 1 and 2 provide a better understanding of EAR and PR respectively by presenting them in a summarized form.

T.S. Babu et al. proposed a swarm-based optimization (PSO) technique to overcome the PSC of a 9 × 9 TCT PV system using EAR. The method was compared with conventional TCT and SDK puzzle PR methods for the power improvement of the PV array [23]. M. Alkhatani et al. discussed a novel repositioning algorithm of PV array based on time-

domain reflectometry for non-uniformly arranged PV arrays [24]. S. Dhanup et al. performed a comparative analysis of a two-phase method with other methods, but the use limits the flexibility of the method in practical applications. EAR techniques require a flexible switching matrix (FSM), a fast-computing processor, data acquisition systems, and fast switching devices that in return increase the overall cost of the system [25].

Table 1. Summary of EAR application on SPV systems.

Ref.	Reconfiguration Technique	Configuration	Control Algorithm	Sensed Parameters	Total Switches
[8]	Dynamic Link Logical Block	SP: 3 × 2	Rough set theory	I	$2(N_{pv}+1)$
[10]	Convert SP to TCT	SP: 3 × 3	Arduino-based DAS	V, I, frequency	$(N_{pv}-2)$
[14]	Switching Matrix	TCT: 9 × 9	Grasshopper Optimization	V, I, Irradiance	$2(i*j)$
[15]	Optimal Switching Matrix	TCT: 9 × 9	Marine predators Algorithm	Irradiance, V	$6(i*j)$
[16]	Switching Matrix	SP: 6 × 5	Refer to [14]	V, I	NS
[23]	Swarm Matrix Dispersion	TCT: 9 × 9	PSO	V, I	2(1 pole m throw)
[24]	Irradiation Equalization	TCT: 3 × 4	Hierarchical Iterative sorting	V, Irradiance	$(i*j)$
[25]	Two-Phase Process	TCT: 9 × 9	Row current Minimization	V, I	2(1 pole m throw)
[26]	Adaptive Bank	TCT	Bubble sort	V, I	$2(i*j)$
[27]	Irradiation Equalization	TCT: 4 × 4	Best worst sorting	V, I	$N_{pv}(i^2-i)$
[28]	Adaptive	6 × 4	Fuzzy estimator	V, I	$i+6$
[29]	Adaptive Bank	TCT: $i \times j$	Scanning algorithm	V, I	$2(i*j)$
[30]	Adaptive Reconfiguration	SP: 3 × 3	Firefly algorithm	V, I	$2(i*j)$

Table 2. Summary of PR application on SPV systems.

Ref.	Reconfiguration Technique	Configuration	Cloud Movement	Complexity	Types of Shade
[7]	SuDoKu	TCT: 9 × 9	No	Moderate	SN, SW, LN, LW
[9]	GA	TCT: 9 × 9	No	High	SW, SN, LN
[11]	Optimal SuDoKu	TCT: 9 × 9	No	High	Discontinuous shade
[17]	Mitigate MMPL	5 × 5	No	Low	SN, SW, LW, LN
[18]	Puzzle shade dispersion	BLTCT: 5 × 4	No	Low	Different patterns
[19]	SuDoKu Puzzle	TCT: 4 × 4	No	Moderate	Different patterns
[20]	SuDoKu Puzzle	6 × 6	No	Moderate	Consistent shading patterns
[21]	SuDoKu Puzzle	Hybrid TCT: 6 × 6	Yes	Moderate	Non-uniform shading
[22]	Lo Shu Technique	TCT: 9 × 9	No	High	SN, SW, LN, LW
[31]	Topsolver Program	TCT: 3 × 3	No	High	SN, LN, LW
[32]	Zig-Zag Arrangement	TCT: 4 × 3	No	Moderate	Corner patterns with horizontal and vertical shading
[33]	Dominance Square	TCT: 5 × 5	No	Moderate	SN, SW, LN, LW
[34]	Magic Square	TCT: 4 × 4	No	Low	Corner patterns with horizontal and vertical shading
[35]	Latin Square	TCT: 4 × 4	No	Moderate	Corner patterns with horizontal and vertical shading

In this paper, a detailed description and modeling of the movement of clouds due to wind velocity and direction are presented. Various possible shadings that occur due to

cloud movement are considered and analyzed. The work focuses on using GWO to disperse the power losses incurred due to different PSC at different times. With the help of the GWO technique, a physical array reconfiguration method is devised using row current minimization as the objective function for the PV array arrangement. Further, a comparative performance analysis of the proposed GWO-BLTCT with four other topologies such as SP, BLTCT, TCT, and SDK-BLTCT is performed for different scenarios like, long narrow (LN), long wide (LW), short narrow (SN), and short wide (SW). The performance indicators used for comparative analysis are power loss, FF, PR, and PE.

2. Description and Modeling of Solar Module

The solar panels are modeled using the single diode model of PV cells as explained in [24]. This model is the most commonly used as it has fewer parameters and is simple to understand. This model simulates the PV array configurations and their performance. Figure 1 represents the single diode model of the PV cell, in this model, the current is generated due to illumination and is known as photocurrent (I_{ph}). The output current (I_{pv}) is represented by (1).

$$I_{pv} = N_p I_{ph} - N_p I_o \left[\exp\left\{ \frac{q\left(V_{pv} + I_{pv} R_s\right)}{N_s A k T} \right\} - 1 \right] \quad (1)$$

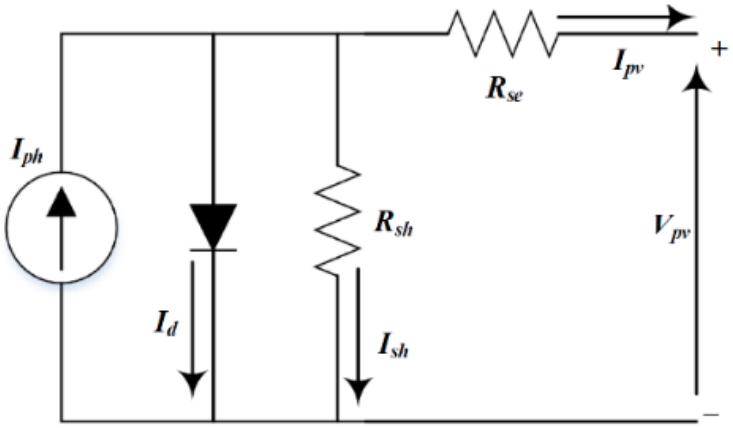


Figure 1. PV cell representing single diode model.

The module considered for the study is TP250MBZ from Tata Power Solar Systems. The in-depth details of the module are provided in Table 3. The study considers a 5 × 5 solar panel system for maximum power extraction. Figure 2 describes the performance of the module at different irradiance and constant temperature of 25 degrees Celsius.

Table 3. Specification of TP250MBZ PV module.

Parameters	Value
Maximum Power	249 W
Open Circuit Voltage (V_{oc})	36.80 V
Voltage at MPP (V_{mpp})	30 V
Cells per Module (N_{cell})	60
Short Circuit Current (I_{sc})	8.83 A
Current at MPP (I_{mpp})	8.30 A
Temperature Coefficient of V_{oc}	−0.33
Temperature Coefficient of I_{sc}	0.063805

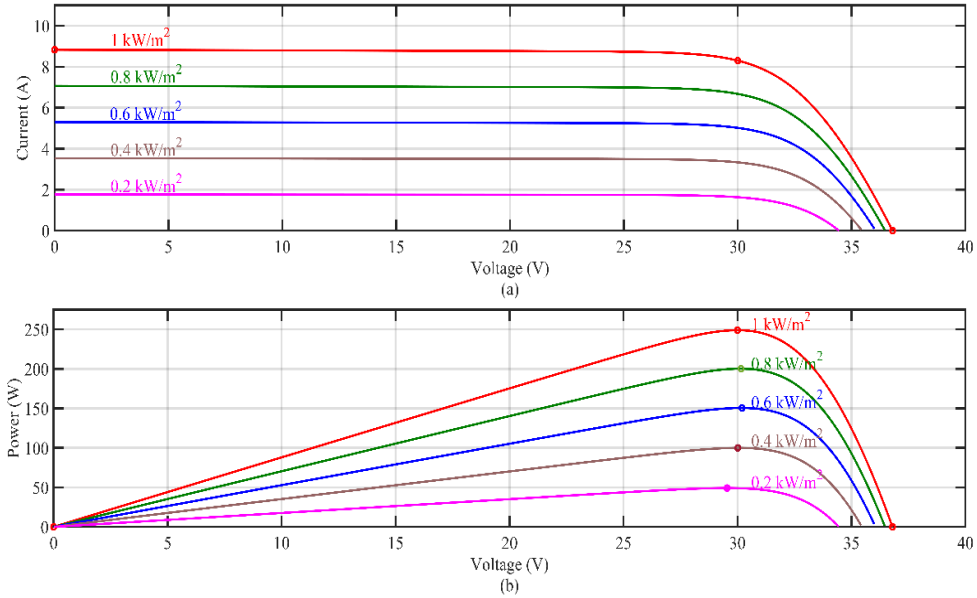


Figure 2. (**a**) I-V characteristic; (**b**) P-V characteristic of TP250MBZ PV module.

3. Cloud Movement and Modeling of Shade

Since the solar PV array is considered as a 5 × 5 matrix, the shade will cause uneven and irregular irradiances on the modules. This reduced irradiance will further impact V_{oc}, I_{sc}. In [36], the authors have modeled the movement of the cloud and explained that the irradiance on the panel is the function of the distance between the center of the module and shadow. In this work, the movement of the cloud is incorporated using acceleration, direction, and wind speed. This movement of cloud creates four different shading conditions, namely, short narrow (SN), long narrow (LN), short wide (SW), and long wide (LW). The authors of [31] have defined the above-mentioned shading conditions based on the width and length of the shading of the modules. The different shading conditions are represented in Figure 3.

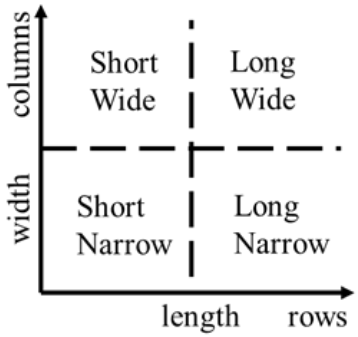


Figure 3. Different shading conditions.

The modeling of various shades due to wind speed and cloud movement is defined using the equations below. Figure 4 shows the various components helpful in modeling partial shade on the solar PV system. Figure 4 represents the variation in cloud movement with time, resulting in non-uniform irradiance on the solar PV module. In this figure, a 5 × 5 solar PV system is considered for which shading conditions are modeled using wind speed, wind direction, the center of the PV module, and the center of the cloud.

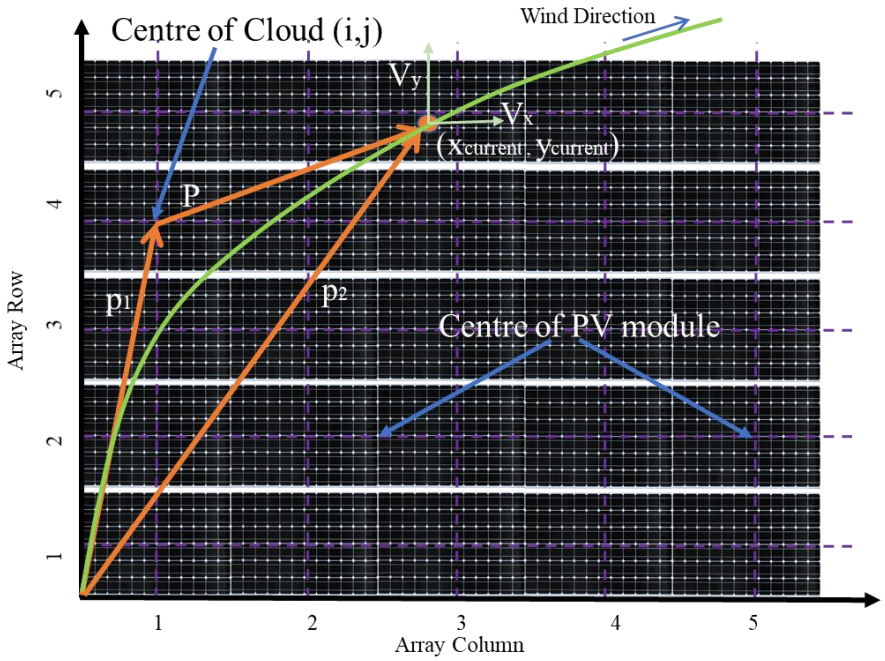


Figure 4. Vector representation of shadow movement on 5 × 5 PV array.

The instantaneous velocity of the center of the shadow is

$$v_{net} = u_{net} + a_{net}t \tag{2}$$

The movement of the clouds can be resolved into vertical and horizontal components, Equation (2) can be modified as

$$v_x = u_x + a_x t \tag{3}$$

$$v_y = u_y + a_y t \tag{4}$$

The vector representation of the center of shadow and PV module at any given time is depicted in Figure 4. Applying the vector addition using triangle law

$$\vec{P}_n = \vec{p}_1 + \vec{p}_2 \tag{5}$$

$$\vec{p}_1 = i\,\hat{\imath} + j\,\hat{\jmath} \tag{6}$$

$$\vec{p}_2 = x_{current}\,\hat{\imath} + y_{current}\,\hat{\jmath} \tag{7}$$

The position of shadow at instance t is

$$x_{current} = x + v_x t \tag{8}$$

$$y_{current} = y + v_y t \tag{9}$$

From Equations (6) and (7), the magnitude of the distance of the shadow from the center of the module is given below

$$\vec{P}_n = (x_{current} - i)\hat{\imath} + (y_{current} - j)\hat{\jmath} \tag{10}$$

$$P_n = \sqrt{(x_{current} - i)^2 + (y_{current} - j)^2} \tag{11}$$

The cloud movement effect on every module is modeled with respect to the change in $I_{sc}(i,j)$ of that particular module [32]. The normalized values of irradiance are given by (12)

$$\frac{I_{sc}(i,j)}{I_{sc}} = 0.55 + \frac{\sin\big(P(i,j)\big)}{1 + P(i,j)} \tag{12}$$

The variation of shade due to the movement of clouds given by the above modeling equations is categorized in four different types of shading conditions as elaborated in Figure 5.

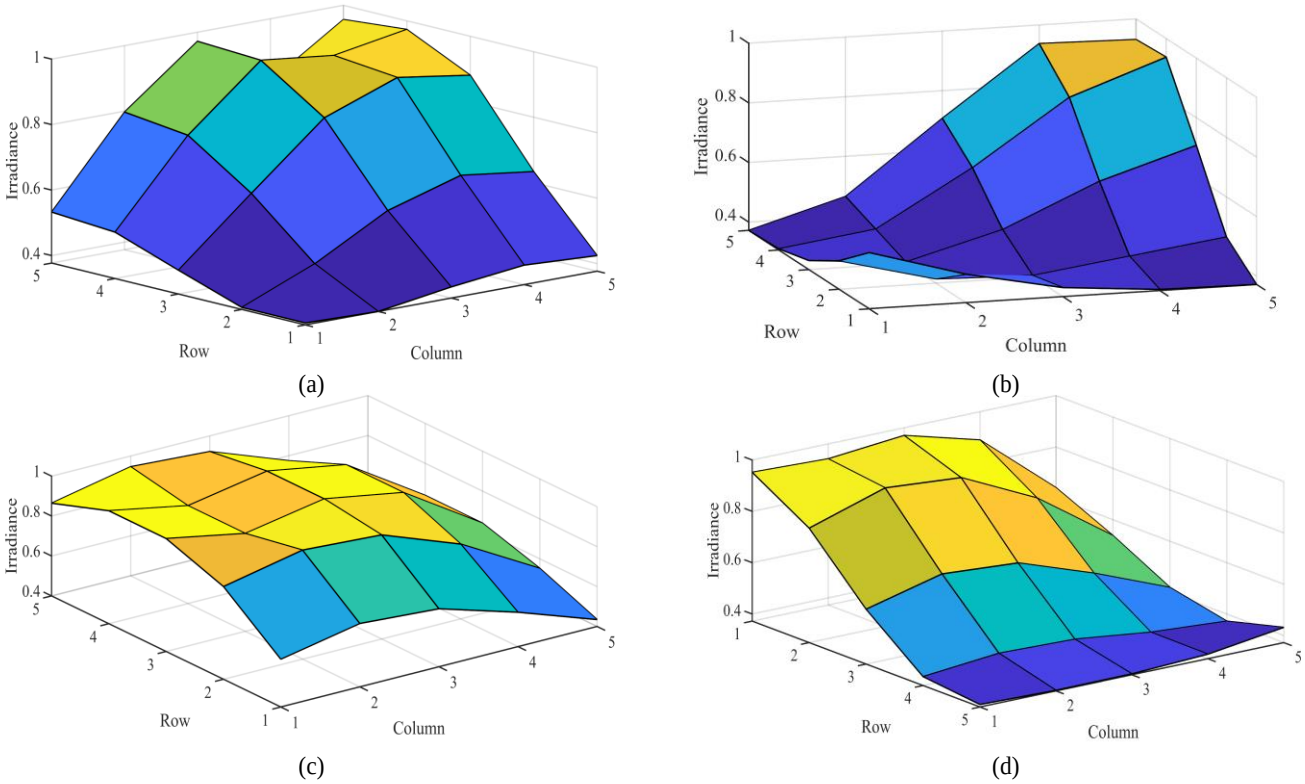


Figure 5. Irregular irradiance at various instances representing different types of shade: (**a**) SW, (**b**) LW, (**c**) SN, and (**d**) LN.

Effect of Partial Shade

The shading of a group of cells results in the reduction of output power. This condition is even more severe when the photodiodes are reversed biased due to partial shading. During this, the diodes offer high resistance and reduce the load current. Furthermore, cells may be damaged due to heat produced as a result of high internal resistance. The output of the solar PV system in the partially shaded condition is different as the P-V characteristics have different MPPs; global and local. The performance indicators considered for the proposed shade dispersion method are mismatch power loss (MMPL), misleading power loss (MLPL), FF, and PR. MMPL occurs because of reduced irradiance on various modules during partial shading conditions, whereas MLPL is the change in power between local MPP and GMPP. Figure 6 represents the different types of power losses, namely mismatch power and misleading power in a PV array during PSC. The fluctuations in P-V characteristics and instability caused due to partial shade can also be mitigated using energy storage systems [37].

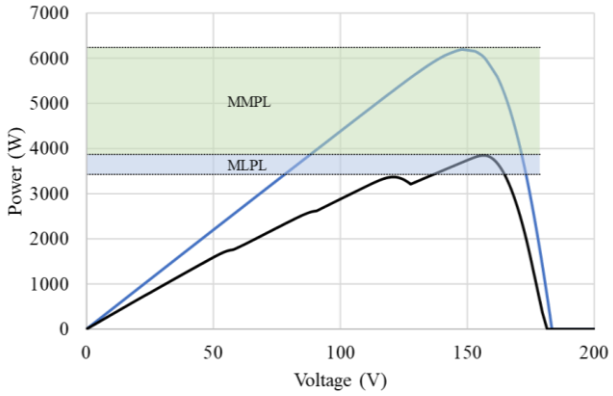


Figure 6. Power loss in the P-V curve during PSC.

$$MMPL = MPP_{wopsc} - GMPP_{psc} \quad (13)$$

$$MLPL = (GMPP - MPP_{local})\ during\ PSC \quad (14)$$

The FF of a solar PV module is the ratio of power at the shaded condition to the product of V_{oc}, I_{sc}. In an ideal condition, this represents a rectangle enclosed within the rectangle formed by V_{oc}, I_{sc}. In the case of a PSC, this shape deviates from a P-V characteristic and is measured in terms of FF.

$$FF = \frac{P_{mx_{psc}}}{V_{oc} I_{sc}} \quad (15)$$

$$PR = \frac{P_{mx_{psc}}}{P_{mx_{stc}}} * 100 \quad (16)$$

4. Proposed Intelligent Model

The application of the intelligent meta-heuristic technique in array reconfiguration is vital because of numerous possibilities of shade dispersion. From the literature, different optimization techniques were available such as particle swarm optimization, ant colony, cuckoo search, and others. However, GWO is a modern nature-inspired optimization technique, capable of solving complex, nonlinear stochastic problems [38]. This algorithm uses the movement of a grey wolf for encircling and targeting its prey [39]. Thus, it does not suffer from low convergence; instead, it is fast and has a low computational burden in comparison. The algorithm uses randomly scattered particles in the search space known as wolves. These wolves are classified into four groups; alphas (α), betas (β), deltas (δ), and omegas (ω). This classification is performed on the basis of hierarchical order, where the alphas are the lead pack of wolves with the least number. The betas follow the decision of the alphas, they are followed by the deltas, and the remaining group of wolves is kept in omegas. The population of wolves increases from alphas to omegas. In optimization, the best solution is given by the alphas (best position) and the entire pack is driven by them. The major benefits of implementing GWO based array reconfiguration are (1) the method effectively converges to the best combination for varying conditions, (2) the algorithm is robust and efficient, (3) the probability of converging to local MPP is reduced, and (4) the speed of convergence is high.

4.1. Proposed Algorithm for Array Reconfiguration

The proposed GWO-based reconfiguration method involves the following steps as given in Figure 7. The hyperparameters required for designing a GWO technique are defined in Table 4. This method consists of four different phases:

- *Initialization Phase*: In this phase initialization of the PV array size, coefficient vectors for the GWO algorithm are done. Where, q is linearly decreased from 2 to 0 over the course of iterations and r_1, r_2 are random vectors in [0, 1].

$$a = 2q.\overrightarrow{r_1} - q \tag{17}$$

$$c = 2.\overrightarrow{r_2} \tag{18}$$

The initial values of wolves are done using (19) and (20) function and, moreover, a pack of them is formed.

$$\overrightarrow{D_{\alpha/\beta/\delta}} = |c.\overrightarrow{X_{\alpha/\beta/\delta}} - \vec{X}(t) \tag{19}$$

$$\overrightarrow{X_{1/2/3}} = \overrightarrow{X_{\alpha/\beta/\delta}} - a.\overrightarrow{D_{\alpha/\beta/\delta}} \tag{20}$$

$\overrightarrow{X_\alpha}, \overrightarrow{X_\beta}, \overrightarrow{X_\delta}$ are the position vectors of α, β, and δ wolves respectively. $\overrightarrow{D_\alpha}$, $\overrightarrow{D_\beta}$, $\overrightarrow{D_\delta}$ define the encircling behavior of α, β, and δ wolves respectively. The equation below defines the updated position vector of the wolf.

$$\overrightarrow{X}(t+1) = \frac{\overrightarrow{X_1} + \overrightarrow{X_2} + \overrightarrow{X_3}}{3} \tag{21}$$

- *Evaluation Phase*: Irradiance (G_{ij}) for every module of the PV array is computed. Here i, j corresponds to row and column numbers. Then, row current (I_{rn}) and array voltage (V_{arr}) is calculated. The objective function for the proposed algorithm is row current difference (D_{rc}) as this should be minimum.

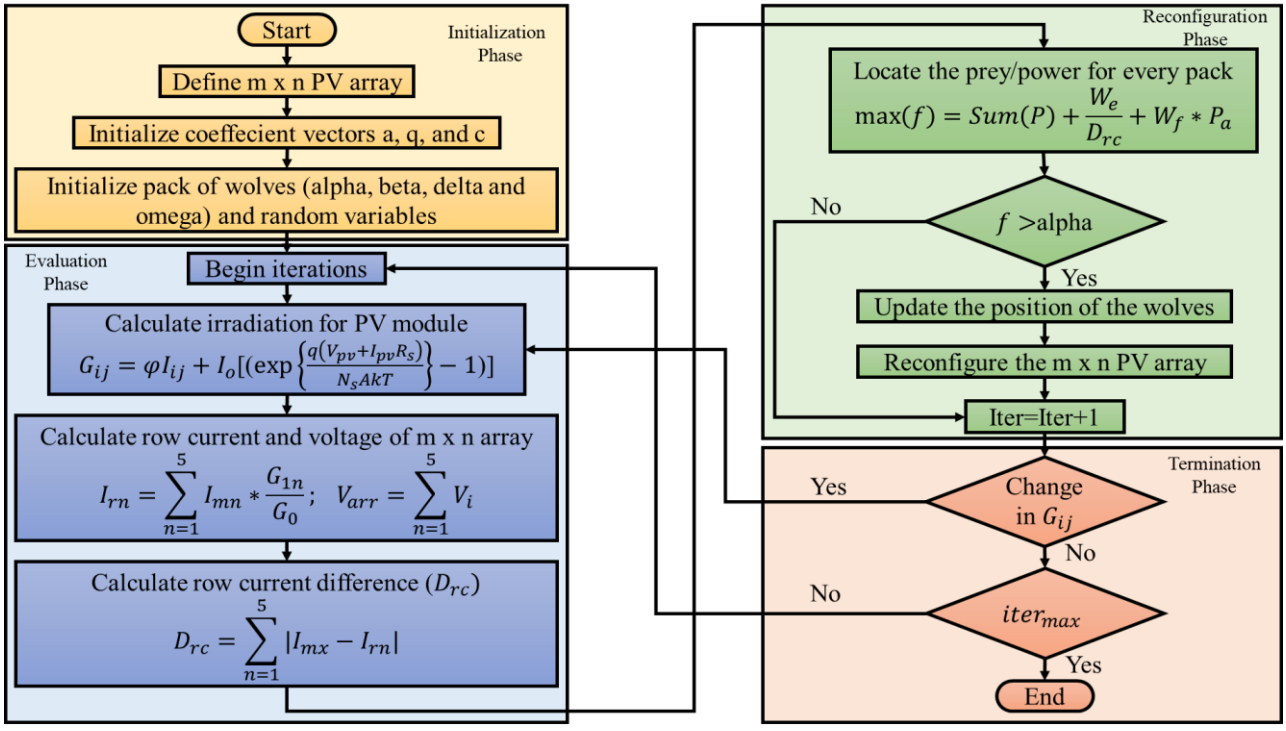


Figure 7. Flowchart for GWO-based array reconfiguration.

- *Reconfiguration Phase*: Based on D_{rc}, I_{rn}, and V_{arr}, the function f has been defined in [23]. Here, P_a is the output power from the PV module. It is to be noted that as the objective function reduces the power output increases of the PV system. By the trial-and-error method the value of W_e and W_f is observed to be 8 and 6 respectively. The updated position of the wolves is calculated by (23).

$$\overrightarrow{D_i} = \left| c \ \overrightarrow{X_p}(t) - \overrightarrow{X}(t) \right| \quad (22)$$

$$\overrightarrow{X_i}(t+1) = \overrightarrow{X_p}(t) - a.\overrightarrow{D} \quad (23)$$

- *Termination Phase:* This stage checks for a change in irradiance (presence of shade) and also evaluated the termination criteria i.e., the maximum iterations. If these conditions are not met then the whole process is reinitiated until the optimal results are obtained.

Table 4. The hyperparameter values for the GWO method.

Parameters	Value
Maximum iterations	1000
q inital	2
q decreasing factor	$2 -$ iteration $* 2/$ max(iterations)
r_1 and r_2	random values between [0, 1]
Tolerance	10^{-6}

4.2. Proposed Grey Wolf Optimizer-Based Bridge-Linked Total Cross-Tied (GWO-BLTCT) Configuration

The GWO algorithm is implemented on the bridge-linked total cross-tied (BLTCT) configuration for the different shading conditions. This results in a physical relocation of modules in optimal pattern resulting in GWO-BLTCT configuration. Figure 8a,b represents the BLTCT and optimal GWO-BLTCT configuration of a 5 × 5 PV array for PSC. Table 5 compares the row currents, voltage and calculates the power for PV array among the BLTCT and proposed GWO-BLTCT type of topologies. The power for each row is calculated by taking the product row current and the potential across that row. Here, in Table 5, I_m, V_m, and P_m denote the maximum current, voltage, and power produced by the module at that respective shading condition.

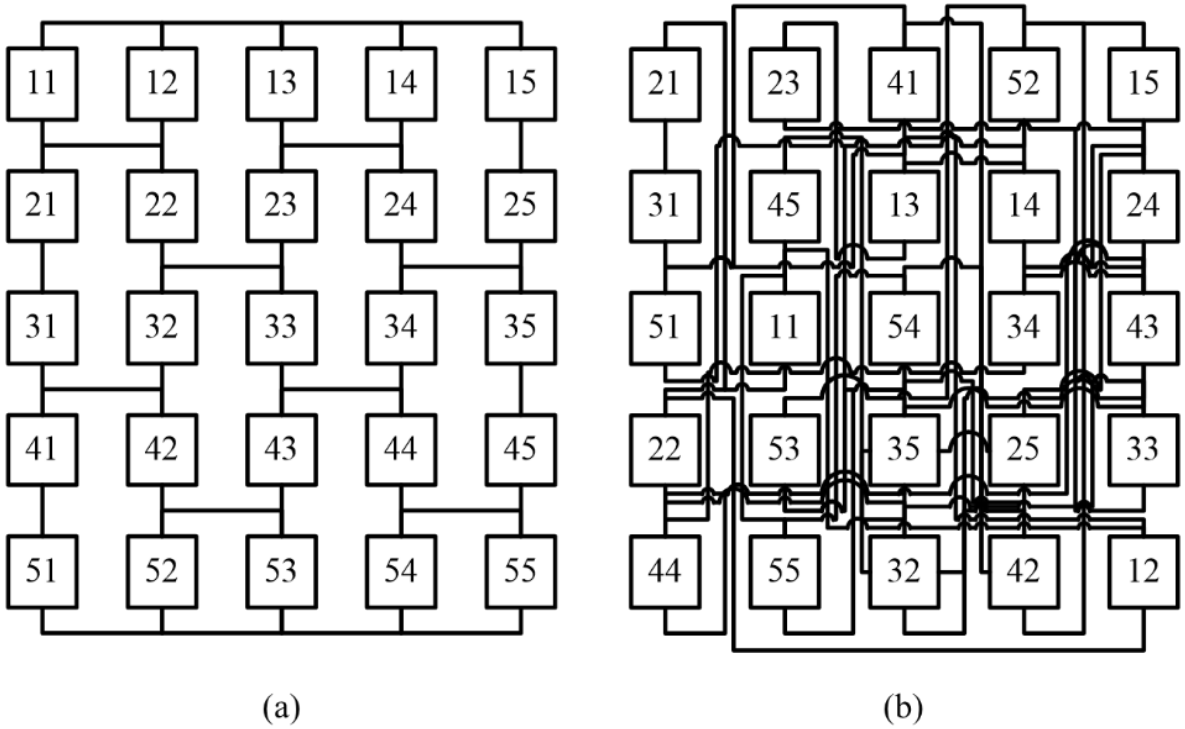


Figure 8. (**a**) BLTCT configuration; (**b**) proposed BLTCT reconfiguration using GWO.

Table 5. Current, voltage, and power calculation of BLTCT and GWO-BLTCT for different shading patterns.

Shading Condition	BLTCT			GWO-BLTCT		
	I_R^{BL} (A)	V (V)	P (W)	I_R^{GWO} (A)	V (V)	P (W)
LN	I_{R5}^{BL} 2.5877 I_m	V_m	2.588 P_m	I_{R5}^{GWO} 2.9837 I_m	V_m	2.984 P_m
	I_{R4}^{BL} 4.1241 I_m	2 V_m	8.248 P_m	I_{R2}^{GWO} 3.0187 I_m	2 V_m	6.037 P_m
	I_{R2}^{BL} 2.7202 I_m	3 V_m	8.161 P_m	I_{R3}^{GWO} 3.1015 I_m	3 V_m	9.305 P_m
	I_{R3}^{BL} 3.4741 I_m	4 V_m	13.896 P_m	I_{R4}^{GWO} 3.1134 I_m	4 V_m	12.454 P_m
	I_{R1}^{BL} 3.0181 I_m	5 V_m	15.091 P_m	I_{R1}^{GWO} 3.2891 I_m	5 V_m	16.446 P_m
LW	I_{R4}^{BL} 2.1795 I_m	V_m	2.180 P_m	I_{R5}^{GWO} 3.2149 I_m	V_m	3.215 P_m
	I_{R5}^{BL} 2.1185 I_m	2 V_m	4.236 P_m	I_{R1}^{GWO} 3.4312 I_m	2 V_m	6.863 P_m
	I_{R2}^{BL} 3.9595 I_m	3 V_m	11.879 P_m	I_{R2}^{GWO} 3.5211 I_m	3 V_m	10.563 P_m
	I_{R1}^{BL} 3.1651 I_m	4 V_m	12.660 P_m	I_{R3}^{GWO} 3.6674 I_m	4 V_m	14.670 P_m
	I_{R3}^{BL} 3.4127 I_m	5 V_m	17.064 P_m	I_{R4}^{GWO} 3.7561 I_m	5 V_m	18.781 P_m
SW	I_{R5}^{BL} 2.5202 I_m	V_m	2.520 P_m	I_{R1}^{GWO} 4.2149 I_m	V_m	4.2149 P_m
	I_{R4}^{BL} 3.6292 I_m	2 V_m	7.258 P_m	I_{R3}^{GWO} 4.4123 I_m	2 V_m	4.4123 P_m
	I_{R3}^{BL} 4.4472 I_m	3 V_m	13.342 P_m	I_{R2}^{GWO} 4.7745 I_m	3 V_m	14.324 P_m
	I_{R2}^{BL} 4.9451 I_m	4 V_m	19.780 P_m	I_{R5}^{GWO} 4.7931 I_m	4 V_m	19.172 P_m
	I_{R1}^{BL} 4.4863 I_m	5 V_m	22.432 P_m	I_{R4}^{GWO} 4.8451 I_m	5 V_m	24.226 P_m
SN	I_{R5}^{BL} 3.0543 I_m	V_m	3.054 P_m	I_{R4}^{GWO} 4.1312 I_m	V_m	4.131 P_m
	I_{R4}^{BL} 4.8588 I_m	2 V_m	9.616 P_m	I_{R1}^{GWO} 4.3171 I_m	2 V_m	8.634 P_m
	I_{R2}^{BL} 4.3513 I_m	3 V_m	13.054 P_m	I_{R2}^{GWO} 4.5078 I_m	3 V_m	13.523 P_m
	I_{R3}^{BL} 4.3513 I_m	4 V_m	17.404 P_m	I_{R3}^{GWO} 4.6122 I_m	4 V_m	18.449 P_m
	I_{R1}^{BL} 4.0588 I_m	5 V_m	20.294 P_m	I_{R5}^{GWO} 4.7409 I_m	5 V_m	23.704 P_m

5. Results and Discussion

The performance of the various configurations namely, SP, TCT, BLTCT, SDK-BLTCT, and GWO-BLTCT is evaluated at four different conditions of partial shade. These shading conditions define the movement of clouds and different types of shadows LN, LW, SW, and SN on the 5 × 5 PV array configuration. The I-V and P-V curves are given for each configuration in Figure 9 and Figure 10 respectively. In both Figures 9 and 10, (a) represents SP, (b) represents TCT, and (c), (d), and (e) depict BLTCT, SDK-BLTCT, and GWO-BLTCT respectively.

The primary objective is to increase the power of the PV array in PSCs by incorporating the movement of clouds. The comparison of achieved GMPP by different PV array configurations in the different PSCs is given in Figure 11a–d. Table 5 shows the global maximum power point of different types of PV arrays in different shading conditions. Table 5 helps in identifying the topology in which maximum power can be extracted for the LN, LW, SW, and SN shading patterns. In LN, SN, and SW partial shade the proposed GWO-BLTCT topology produces the best global maximum power point. Furthermore, during LW shading GWO-BLTCT is better than SDK-BLTCT but the difference of 10.42 W is nominal and can be ignored when considering the significant advantage in GMPP in LN, SN, and SW shading patterns. From Table 6 and Figure 11c,d it is evident that the performance of TCT and BLTCT is closely matched in SW and SN types of shades.

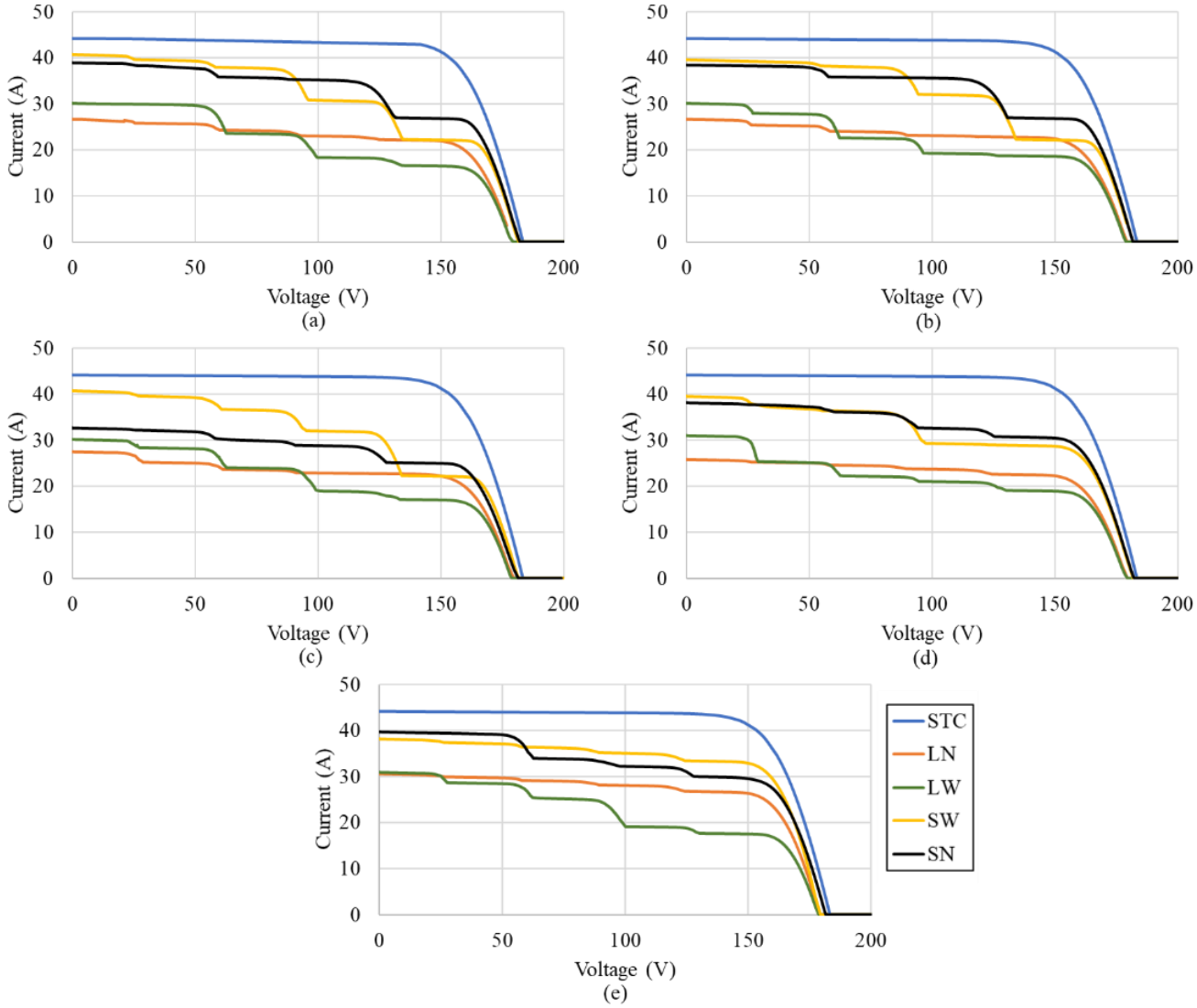


Figure 9. I-V curves at different shading conditions for various configurations: (**a**) SP, (**b**) TCT, (**c**) BLTCT, (**d**) SDK BLTCT, and (**e**) GWO BLTCT.

Table 6. GMPP at different shading conditions at 25 °C.

Configurations	Global Maximum Power Point (W)			
	LN	LW	SW	SN
SP	3308.71	2563.16	3754.20	3848.51
TCT	3367.33	2864.67	3873.38	4208.90
BLTCT	3387.37	2629.98	3865.39	4219.19
SDK-BLTCT	3727.01	2895.45	4572.25	4634.51
GWO-BLTCT	3971.68	2905.87	4960.41	4977.97

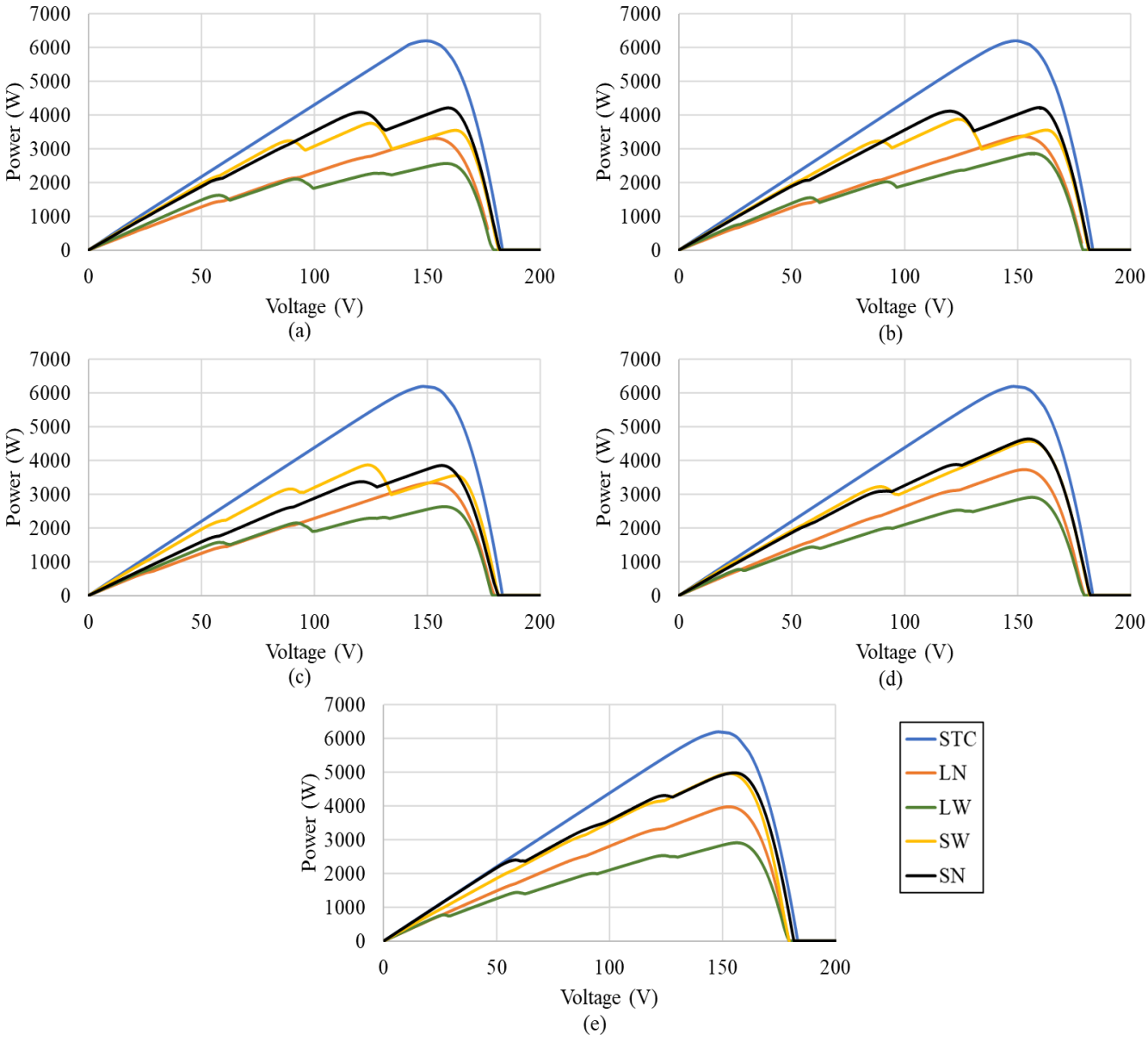


Figure 10. P-V curves at different shading conditions for various configurations: (**a**) SP, (**b**) TCT, (**c**) BLTCT, (**d**) SDK BLTCT, and (**e**) GWO BLTCT.

Power enhancement (PE) is calculated for all the topologies with respect to the SP configurations and is calculated using (23)

$$PE = \frac{GMPP_{TCT|BLTCT|SDK|GWO} - GMPP_{SP}}{GMPP_{SP}} * 100 \quad (24)$$

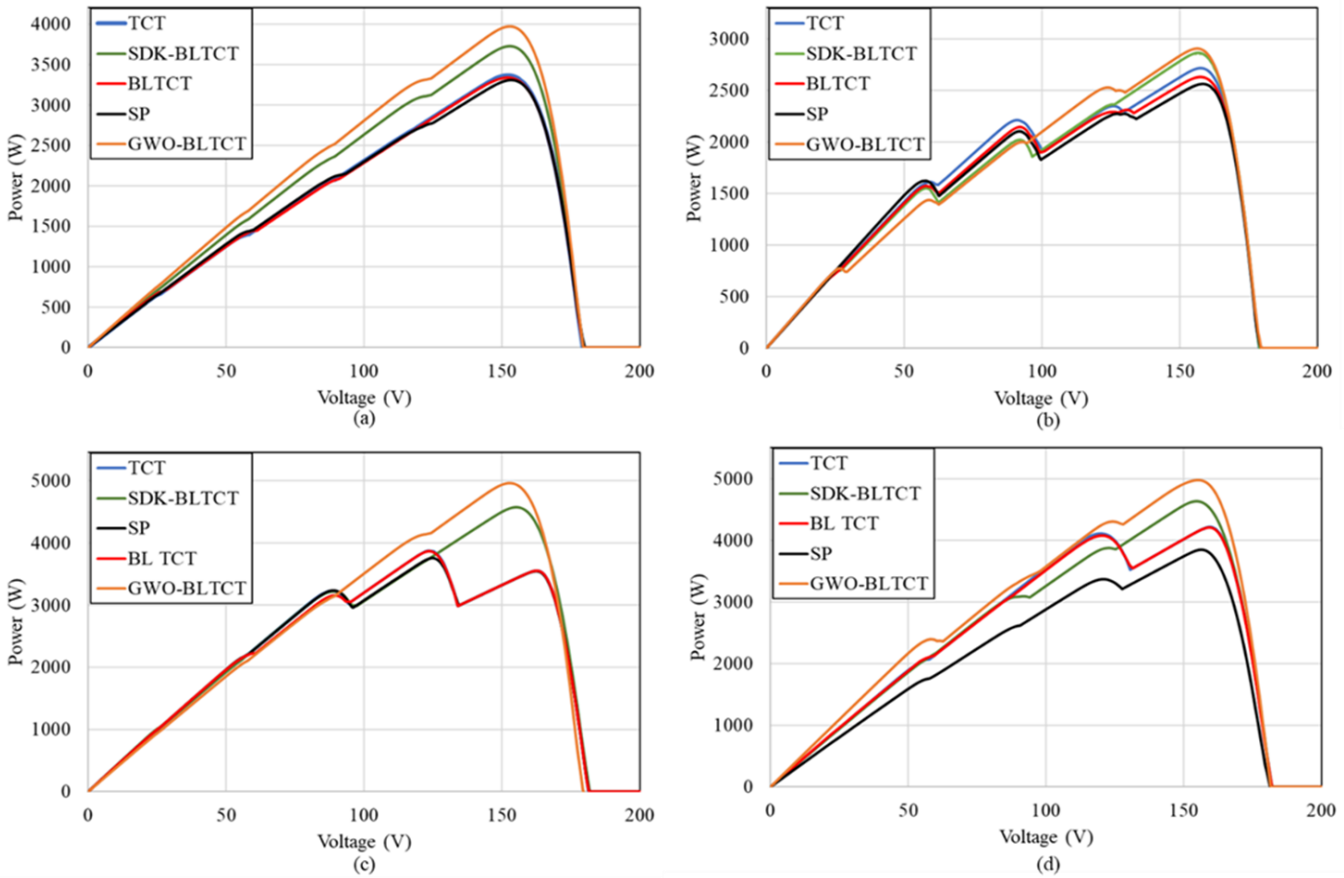


Figure 11. Effect of different shade due to movement of cloud on the PV array: (**a**) LN (**b**) LW (**c**) SW and (**d**) SN.

Here, $GMPP_{SP}$ is the global maximum power point achieved by SP configuration, and $GMPP_{TCT|BLTCT|SDK|GWO}$ is the global maximum power point achieved by TCT, BLTCT, SDK-BLTCT, or GWO-BLTCT configurations. From Table 7, it is observed that the LW type of shade due to wind speed and cloud movement is the most challenging type of shade as the performance of all the topologies is most hampered during it. The best performing topologies during this instance is proposed GWO-BLTCT and marginally behind is SDK-BLTCT. The performance ratio is also the least (<50%) for the LW shade for all the topologies, whereas for all the other types of shades LN, SW, and SN, the performance ratio is high, i.e., more than 50%. Among them, the SN type of shade is the least shaded. The proposed GWO-BLTCT type of configuration is the best performing configuration for LN, SN, and SW types of shade. Power enhancement of 20.11%, 32.14%, and 29.37% is achieved for LN, SN, and SW respectively, being the highest among the rest of the topologies. Figure 12 represents the performance indicators of all the above topologies for different types of shades. Figure 12a shows power loss, Figure 12b depicts FF, Figure 12c,d plots PR and PE respectively for the four shading conditions.

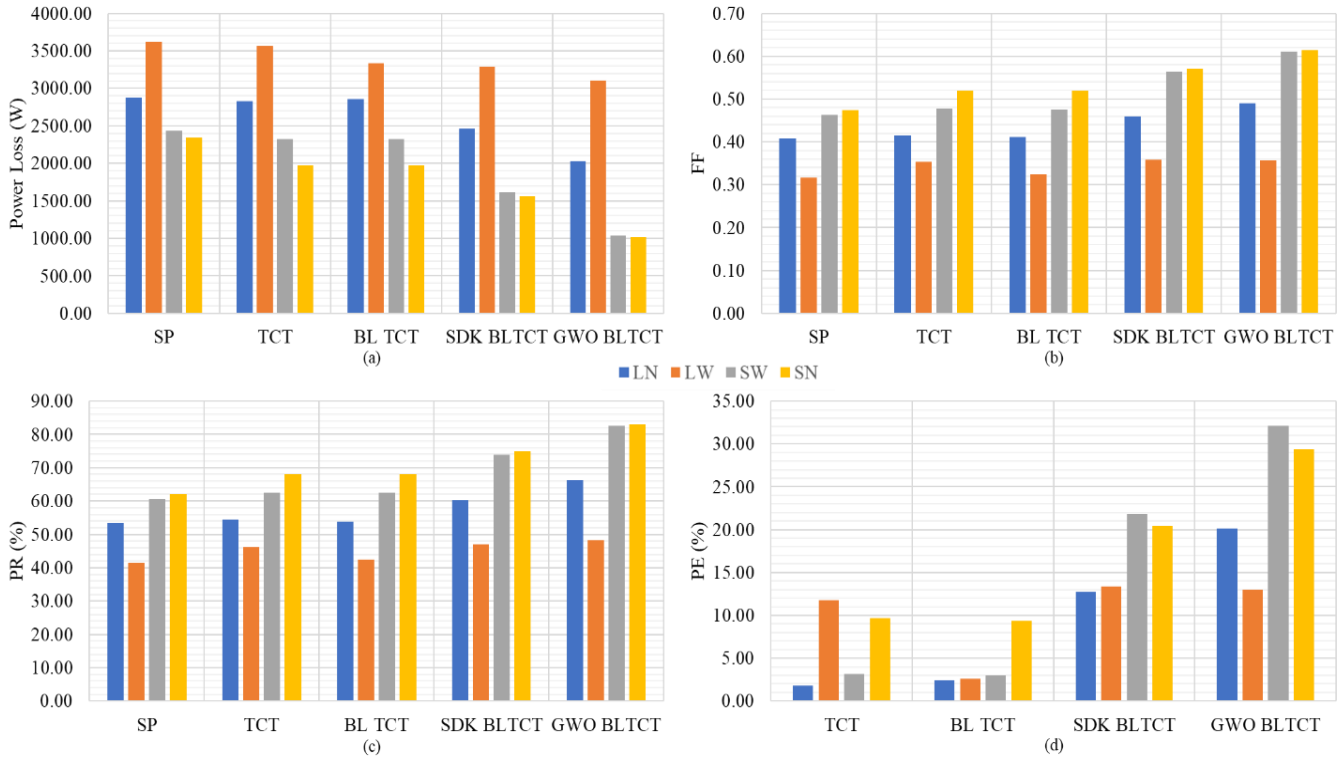


Figure 12. Performance indicators: (**a**) power Loss, (**b**) fill-factor, (**c**) PR, and (**d**) PE.

Table 7. Performance of the PV array in partial shading conditions due to cloud movement at different instances.

Configuration	Performance Indicator	LN	LW	SW	SN
SP	MMPL (W)	2875.66	3621.21	2430.17	2343.43
	PR (%)	53.50	41.45	60.71	62.16
	FF	0.41	0.32	0.46	0.47
	PE (%)	Nil	Nil	Nil	Nil
TCT	MMPL (W)	2827.50	3561.96	2321.45	1975.64
	PR (%)	54.36	46.25	62.53	68.12
	FF	0.41	0.35	0.48	0.52
	PE (%)	1.80	11.77	3.18	9.65
BLTCT	MMPL (W)	2854.57	3330.16	2326.55	1975.47
	PR (%)	53.91	42.48	62.44	68.06
	FF	0.41	0.32	0.48	0.52
	PE (%)	2.41	2.61	2.97	9.38
SDK-BLTCT	MMPL (W)	2464.93	3104.55	1619.69	1557.43
	PR (%)	60.20	46.94	73.85	74.86
	FF	0.46	0.36	0.56	0.57
	PE (%)	12.70	12.97	21.80	20.44
GWO-BLTCT	MMPL (W)	2028.32	3286.07	1039.59	1022.03
	PR (%)	66.19	48.26	82.67	82.97
	FF	0.49	0.36	0.61	0.61
	PE (%)	20.11	13.38	32.14	29.37

6. Conclusions

The performance of the PV system is affected adversely by various factors such as solar irradiance, cell temperature, air mass coefficient, and wind speed. In this paper impact of partial shading on a 5 × 5 PV system is discussed in detail. The paper proposed a physical array reconfiguration technique based on the GWO method for a 5 × 5 PV array arrangement for different types of partial shading. In this work partial shade is modeled by the movement of the center of the cloud, wind speed, and wind direction. The paper models the realistic movement of could considering wind direction and wind velocity at different time instances. This movement of the cloud induces partial shade onto the PV system which is classified as LW, LN, SN, and SW types of shade. The proposed GWO algorithm uses row current minimization for solving the objective function and reconfigured the position of PV panels in a BLTCT type configuration. The performance of the proposed GWO-BLTCT was analyzed using performance indicators, such as FF, power loss, PR, and PE. The proposed GWO-BLTCT was compared with other topologies/configurations, SP, TCT, BLTCT, and SDK-BLTCT for four different types of shades: LN, LW, SW, and SN. For the proposed GWO-BLTCT array configuration GMPP, PR, PE, and FF were the highest in comparison to other configurations; moreover, the loss component was minimum. Overall, for the proposed configuration the average PE and PR are observed to be 23.75% and 70.02% respectively. Thus, the GWO-BLTCT PV array configuration is found to be superior when compared with SP, TCT, BLTCT, and SDK-BLTCT.

Author Contributions: Conceptualization, A.K. and M.R.; methodology, A.K. and U.N.; software, A.K.; validation, M.R. and M.A.; formal analysis, U.N. and M.A.; investigation, A.K.; resources, M.A.; data curation, M.A.; writing—original draft preparation, A.K.; writing—review and editing, M.R. and U.N.; visualization, A.K. and M.A.; supervision, M.R. and U.N.; project administration, M.A.; funding acquisition, M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare there is no conflict of interest.

Nomenclature and Abbreviations

I_{pv}	PV output current
N_p	Number of parallel cells
I_{ph}	Photon current
I_o	Diode saturation current
q	Electron charge
V_{pv}	Voltage of PV module
R_s	Series resistance
N_s	Number of series cells
A	Ideality factor
k	Boltzman constant
T	Actual module operating temperature in K
v_{net}	Velocity at center
u_{net}	Initial velocity of wind
a_{net}	Instantaneous acceleration
$\vec{P}_n$	Net position vector
$\overrightarrow{D_{\alpha/\beta/\delta}}$	Distance vector of $\alpha/\beta/\delta$ wolves from prey
$\overrightarrow{X_{\alpha/\beta/\delta}}$	Position vector of $\alpha/\beta/\delta$ wolves
G_{ij}	Irradiance incident on (i^{th},j^{th}) PV panel of array
P_a	PV module power output
D_{rc}	Difference of row current
V_{arr}	Array voltage
I_{rn}	n^{th} Row current
$P_{mx_{psc}}$	Maximum power at PSC

$P_{mx_{stc}}$	Maximum power at STC
W_e and W_f	Assumed weights
BLTCT	Bridge-linked total cross-tied
FF	Fill-factor
$GMPP_{psc}$	Global MPP during PSC
$GMPP_{SP}$	Global MPP of series-parallel
GWO BLTCT	Grey wolf optimizer BLTCT
MMPL	Mismatched power loss
MLPL	Misleading power loss
MPP_{wopsc}	Maximum power point without PSC
PS	Partial shading
PE	Power enhancement
PR	Performance ratio
PSC	Partial shading condition
SDK BLTCT	SuDoKu bridge linked TCT
STC	Standard testing conditions
SP	Series-parallel
TCT	Total cross-tied

References

1. World Energy Trilemma Index 2020. World Energy Council UK. Available online: https://trilemma.worldenergy.org/reports/main/2020/ (accessed on 18 March 2021).
2. Ong, S.; Campbell, C.; Denholm, P.; Margolis, R.; Heath, G. Land-Use Requirements for Solar Power Plants in the United States. Available online: http://www.nrel.gov/docs/fy13osti/56290.pdf (accessed on 11 April 2021).
3. Green, M.A.; Hishikawa, Y.; Warta, W.; Dunlop, E.D.; Levi, D.H.; Hohl-Ebinger, J.; Ho-Baillie, A.W. Solar cell efficiency tables (version 50). *Prog. Photovolt. Res. Appl.* **2017**, *25*, 668–676.
4. Ahmed, J.; Salam, Z. A critical evaluation on maximum power point tracking methods for partial shading in PV systems. *Renew. Sustain. Energy Rev.* **2015**, *47*, 933–953.
5. Ansari, S.; Ayob, A.; Lipu, M.S.H.; Saad, M.H.M.; Hussain, A. A Review of Monitoring Technologies for Solar PV Systems Using Data Processing Modules and Transmission Protocols: Progress, Challenges and Prospects. *Sustainability* **2021**, *13*, 8120–8154.
6. Kumar, A.; Thapar, S. Addressing Land Issues for Utility Scale Renewable Energy Deployment in India. Teri School of Advance Studies, India. Available online: https://shaktifoundation.in/wp-content/uploads/2018/01/Study-Report-Addressing-Land-Issues-for-Utility-Scale-Renewable-Energy-Deployment-in-India.pdf (accessed on 20 June 2021).
7. Rani, B.I.; Ilango, G.S.; Nagamani, C. Enhanced power generation from PV array under partial shading conditions by shade dispersion using Su Do Ku configuration. *IEEE Trans. Sustain. Energy* **2013**, *4*, 594–601.
8. Vicente, P.D.S.; Pimenta, T.C.; Ribeiro, E.R. Photovoltaic Array Reconfiguration Strategy for Maximization of Energy Production. *Int. J. Photoenergy* **2015**, *3*, 592383. https://doi.org/10.1155/2015/592383.
9. Deshkar, S.N.; Dhale, S.B.; Mukherjee, J.S.; Babu, T.S.; Rajasekar, N.; Solar, P.V. array reconfiguration under partial shading conditions for maximum power extraction using genetic algorithm. *Renew. Sustain. Energy Rev.* **2015**, *43*, 102–110.
10. Pachauri, R.; Singh, R.; Gehlot, A.; Samakaria, R.; Choudhury, S. Experimental analysis to extract maximum power from PV array reconfiguration under partial shading conditions. *Eng. Sci. Technol. Int. J.* **2019**, *22*, 109–130.
11. Potnuru, S.R.; Pattabiraman, D.; Ganesan, S.I.; Chilakapati, N. Positioning of PV panels for reduction in line losses and mismatch losses in PV array. *Renew. Energy* **2015**, *78*, 264–275.
12. Eltamaly, A.M.; Farh, H.M.H. Dynamic global maximum power point tracking of the PV systems under variant partial shading using hybrid GWO-FLC. *Sol. Energy* **2019**, *177*, 306–316.
13. Camilo, J.C.; Guedes, T.; Fernandes, D.A.; Melo, J.D.; Costa, F.F.; Sguarezi, A.J. A maximum power point tracking for photovoltaic systems based on Monod equation. *Renew. Energy* **2019**, *130*, 428–438.
14. Fathy, A. Recent meta-heuristic grasshopper optimization algorithm for optimal reconfiguration of partially shaded PV array. *Sol. Energy* **2018**, *171*, 638–651.
15. Yousri, D.; Babu, T.S.; Beshr, E.; Eteiba, M.B.; Allam, D. A Robust Strategy Based on Marine Predators Algorithm for Large Scale Photovoltaic Array Reconfiguration to Mitigate the Partial Shading Effect on the Performance of PV System. *IEEE Access* **2020**, *8*, 112407–112426.
16. Liu, X.; Wang, Y. Reconfiguration method to extract more power from partially shaded photovoltaic arrays with series-parallel topology. *Energies* **2019**, *12*, 1439–1454.
17. Pendem, S.R.; Mikkili, S. Modelling and performance assessment of PV array topologies under partial shading conditions to mitigate the mismatching power losses. *Sol. Energy* **2018**, *160*, 303–321.
18. Yadav, A.S.; Pachauri, R.K.; Chauhan, Y.K. Comprehensive investigation of PV arrays with puzzle shade dispersion for improved performance. *Sol. Energy* **2016**, *129*, 256–285.

19. Yadav, A.S.; Pachauri, R.K.; Chauhan, Y.K. Comprehensive investigation of PV arrays under different shading patterns by shade dispersion using puzzled pattern based Su-Do-Ku puzzle configuration. In Proceedings of the 2015 1st International Conference on Next Generation Computing Technologies, Dehradun, India, 4–5 September 2015; Volume 1, pp. 824–830.
20. Vijayalekshmy, S.; Bindu, G.R.; Rama, I.S. A novel Zig-Zag scheme for power enhancement of partially shaded solar arrays. *Sol. Energy* **2016**, *135*, 92–102.
21. Sagar, G.; Pathak, D.; Gaur, P.; Jain, V. A Su Do Ku puzzle based shade dispersion for maximum power enhancement of partially shaded hybrid bridge-link-total-cross-tied PV array. *Sol. Energy* **2020**, *204*, 161–180.
22. Venkateswari, R.; Rajasekar, N. Power enhancement of PV system via physical array reconfiguration based Lo Shu technique. *Energy Convers. Manag.* **2020**, *215*, 112885.
23. Babu, T.S.; Ram, J.P.; Dragičević,, T.; Miyatake, M.; Blaabjerg, F.; Rajasekar, N. Particle swarm optimization based solar PV array reconfiguration of the maximum power extraction under partial shading conditions. *IEEE Trans. Sustain. Energy* **2018**, *9*, 74–85.
24. Alkahtani, M.; Wu, Z.; Kuka, C.S.; Alahammad, M.S.; Ni, K. A Novel PV Array Reconfiguration Algorithm Approach to Optimising Power Generation across Non-Uniformly Aged PV Arrays by Merely Repositioning. *J. Multidiscip. Sci.* **2020**, *3*, 32–53.
25. Pillai, D.S.; Rajasekar, N.; Ram, J.P.; Chinnaiyan, V.K. Design and testing of two phase array reconfiguration procedure for maximizing power in solar PV systems under partial shade conditions (PSC). *Energy Convers. Manag.* **2018**, *178*, 92–110.
26. Mehta, H.K.; Warke, H.; Kukadiya, K.; Panchal, A.K. Accurate Expressions for Single-Diode-Model Solar Cell Parameterization. *IEEE J. Photovolt.* **2019**, *9*, 803–810.
27. Storey, J.P.; Wilson, P.R.; Bagnall, D. Improved optimization strategy for irradiance equalization in dynamic photovoltaic arrays. *IEEE Trans. Power Electron.* **2013**, *28*, 2946–2956.
28. Tabanjat, A.; Becherif, M.; Hissel, D. Reconfiguration solution for shaded PV panels using switching control. *Renew. Energy* **2015**, *82*, 4–13.
29. Parlak, K.Ş. PV array reconfiguration method under partial shading conditions. *Int. J. Electr. Power Energy Syst*. **2014**, *63*, 713–721.
30. Nazeri, M.N.R.; Tajuddin, M.F.N.; Babu, T.S.; Azmi, A.; Malvoni, M.; Kumar, N.M. Firefly Algorithm-Based Photovoltaic Array Reconfiguration for Maximum Power Extraction during Mismatch Conditions. *Sustainability* **2021**, *13*, 3206.
31. Villa, L.F.L.; Picault, D.; Raison, B.; Bacha, S.; Labonne, A. Maximizing the power output of partially shaded photovoltaic plants through optimization of the interconnections among its modules. *IEEE J. Photovolt.* **2012**, *2*, 154–163.
32. Vijayalekshmy, S.; Bindu, G.R.; Iyer, S.R. Performance of partially shaded photovoltaic array configurations under shade dispersion. In Proceedings of the 2014 International Conference on Advances in Green Energy, Thiruvananthapuram, India, 17–18 December 2014; pp. 1–6.
33. Dhanalakshmi, B.; Rajasekar, N. Dominance square based array reconfiguration scheme for power loss reduction in solar Photovoltaic (PV) systems. *Energy Convers. Manag.* **2018**, *156*, 84–102.
34. Yadav, A.S.; Pachauri, R.K.; Chauhan, Y.K.; Choudhury, S.; Singh, R. Performance enhancement of partially shaded PV array using novel shade dispersion effect on magic-square puzzle configuration. *Sol. Energy* **2017**, *144*, 780–797.
35. Sahu, H.S.; Nayak, S.K.; Mishra, S. Maximizing the Power Generation of a Partially Shaded PV Array. *IEEE J. Emerg. Sel. Top. Power Electron.* **2016**, *4*, 626–637.
36. Nguyen, D.D.; Lehman, B. Modeling and simulation of solar PV arrays under changing illumination conditions. In Proceedings of the 2006 IEEE Workshop on Computers in Power Electronics, Troy, NY, USA, 16–19 July 2006; pp. 295–299.
37. Andrychowicz, M. The Impact of Energy Storage along with the Allocation of RES on the Reduction of Energy Costs Using MILP. *Energies* **2021**, *14*, 3783–3798.
38. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey Wolf Optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61.
39. Kumar, A.; Rizwan, M.; Nangia, U. A Hybrid Intelligent Approach for Solar Photovoltaic Power Forecasting: Impact of Aerosol Data. *Arab. J. Sci. Eng*. **2020**, *45*, 1715–1732.

View publication stats

Front. Comput. Sci., 2022, 16(4): 164205
https://doi.org/10.1007/s11704-021-0127-0

RESEARCH ARTICLE

Handling class imbalance problem in software maintainability prediction: an empirical investigation

Ruchika MALHOTRA[1], **Kusum LATA** (✉)[1,2]

1 Discipline of Software Engineering, Department of Computer Science & Engineering, Delhi Technological University, Delhi 110042, India
2 University School of Management and Entrepreneurship, East Delhi Campus, Delhi Technological University, Delhi 110095, India

© Higher Education Press 2022

Abstract As the complexity of software systems is increasing; software maintenance is becoming a challenge for software practitioners. The prediction of classes that require high maintainability effort is of utmost necessity to develop cost-effective and high-quality software. In research of software engineering predictive modeling, various software maintainability prediction (SMP) models are evolved to forecast maintainability. To develop a maintainability prediction model, software practitioners may come across situations in which classes or modules requiring high maintainability effort are far less than those requiring low maintainability effort. This condition gives rise to a class imbalance problem (CIP). In this situation, the minority classes' prediction, i.e., the classes demanding high maintainability effort, is a challenge. Therefore, in this direction, this study investigates three techniques for handling the CIP on ten open-source software to predict software maintainability. This empirical investigation supports the use of resampling with replacement technique (RR) for treating CIP and develop useful models for SMP.

Keywords software maintenance, software maintainability, imbalanced learning

1 Introduction

After the initial release, software systems are subjected to continuous modifications and evolution to respond to the changes in the real world. The changes in the software systems are due to multiple causes, few of which include enhancing existing features, adding new features, and correction of the existing defects. As the complexity of software is increasing, maintaining such complex systems is challenging of software practitioners. Researchers have developed many SMP models to estimate the maintainability early during the software development [1−3]. Indeed, early prediction of software maintainability supports to distribute the limited resources of an organization optimally. Also, classes demanding high maintainability effort can be redesigned better, thus reducing the overall cost and time spent on the software product. Predicting software maintainability is complex as software maintenance is targeted on various kinds of interventions like adaptations and corrections. Over the year's researchers agreed that SMP models could be constructed by forming a relationship between maintainability and software metrics. The maintainability is quantified in terms of changes made in the code during the maintenance phase of the software development life cycle [4,5]. The software metrics define different software characteristics like size, complexity, inheritance depth, cohesion, and coupling. The studies in the literature applied various Statistical and Machine Learning (ML) techniques to forecast software maintainability of proprietary and open-source software datasets [6]. The process of maintainability prediction model development requires a dataset for training the model. The dataset should comprise an adequate quantity of datapoints of high maintainability effort and low maintainability effort classes to properly train the model to classify the cases of each class with reasonable accuracy. Unfortunately, in the maintenance projects, the majority of classes require more straightforward interventions to a few classes requiring high or complex interventions, resulting in an imbalanced dataset. If the model is trained using such an imbalanced dataset, the probability of a minority class's misclassification increases because the prediction model may not have learned the sufficient number of data points of the minority class well. Such prediction models do not have any practical relevance as they fail to predict a minority class (i.e., class requiring high maintainability effort) accurately. The consequence of such misclassification would lead to inappropriate allocation of resources to the misclassified classes. Therefore, this may result in poor quality software products.

Therefore, the situations where we have the imbalanced datasets, techniques to handle such datasets are essential for adequately training the prediction models. Thus, in this direction, the study explores three data resampling methods, namely: Synthetic Minority Oversampling Technique (SM),

Received March 28, 2020; accepted February 5, 2021

E-mail: kusumlata@dtu.ac.in

Spread Subsampling (SS), and Resample with replacement (RR) to deal with imbalanced data. The contributions of the study are enumerated as follows:

- The study uses various software metrics to describe software characteristics for developing SMP models, some of which have not been used before in predicting software maintainability.
- We present an effective investigation to observe the effect of data resampling methods to improve SMP models' performance in terms of predicting low and high maintainability effort classes correctly.
- The study makes use of appropriate performance measures to assess and analyse the predictive performance of SMP models.
- A meticulous statistical analysis to support the results by statistical tests is carried out in the study.

The following research questions-RQs are inspected in this study.
RQ1) How are the predictions of SMP models built with datasets that are imbalanced?
RQ2) Do the data resampling methods results in improved prediction models?
RQ3) Among the data resampling methods employed, which performs best to handle the imbalanced data?

All of the above research questions are evaluated using ten open-source data sets. We built SMP models using ML techniques. Ten-fold cross-validation (TFCV) is used to split the data sets into train-test subsets. The predictive performance of the SMP models is assessed by AUC (Area under Curve) and G-Mean. We also use precision, accuracy, specificity, and sensitivity for evaluating SMP models' performance. The study employs statistical tests to compare the different methods to handle imbalanced data. The results support the RR technique to handle CIP amongst all other data resampling methods used in this paper. The paper is divided into eight sections: Section 2 presents the summary of related work, and Section 3 expounds the design elements. In Section 4, we present the experimental framework. Section 5 puts forward the research methodology. In Section 6, experimental results are documented. Section 7 briefs the threats to the validity of our study. Section 8 presents discussion. The conclusions and future work are presented in Section 9.

2 Related work

The related work is presented in three sections. The work done in literature for SMP is outlined in the first section. The second section presents the studies that contributed to handle CIP, and the third section exhibits the studies that dealt with CIP in the software engineering domain.

2.1 Software maintainability prediction studies

In literature, various researchers have successfully developed SMP models with the by applying numerous kinds of techniques. In the literature, the relationship of software metrics with maintainability is well proven. The software metrics quantitively characterize a software. In the paper by Thwin and Quah [7], the relation between object-oriented (OO) metrics and maintainability is established. Two types of neural networks, namely general regression neural network and ward network, are used to construct SMP models. The papers by Van Koten and Gray [8] and Aggarwal et al. [9] used OO metrics pertaining to two commercial software projects to develop SMP models. Kaur and Kaur validated two OO systems using 27 ML and statistical methods. The datasets from software projects implemented in C++ language are used by Zhang et al. [10] to construct models to predict maintainability of software classes. The study used 44 software metrics. The metrics are collected at different levels of granularity. Malhotra and Chug have given a benchmarking framework for maintainability prediction, where the study uses OO metrics collected from seven open-source software systems written in Java language. Recently, Malhotra and Lata [3], Kumar and Rath [11] evaluated various OO metrics to ascertain the relationship with software maintainability. These studies reflect the researchers' efforts to develop SMP models and confirm that software matrices play an essential role in building these models. However, developing SMP models in the case of imbalanced datasets is still a challenge. At the time of building SMP models, CIP comes up when the classes with high manageability effort are deficient compared to low maintainability effort classes.

2.2 Approaches to handle CIP

Numerous techniques have been found in the literature to handle the CIP. This section gives a brief of the state-of-the-art to deal with CIP. From the literature, two broad ways to deal with this problem have been found. These ways are data resampling and cost-sensitive learning. In the data resampling methods, imbalanced sample space is balanced with oversampling and undersampling. These methods modify the dataset in such a way that after resampling, the dataset contains an adequate quantity of samples of the majority and minority class. Undersampling methods tend to make a balance between the majority class and minority class by randomly discarding the data points of the majority class. In contrast, oversampling methods make a balance between the two classes by replicating the instances of the minority class. The cost-sensitive learning algorithms tend to assign higher misclassification cost to the minority class data samples than the majority class samples. The studies employing different approaches to handle imbalanced learning are presented in this section. Zhou [12] investigated different resampling methods and suggested different sample ratios for different resampling approaches. The studies Napierian and Stefanowski [13] and Loyola-González et al. [14] studied the performance of diverse resampling methods. Essential insights from these studies are as follows: (i) for a dataset consisting of hundreds of instances undersampling methods are superior as compare to oversampling methods regarding computational time; (ii) oversampling method synthetic minority oversampling technique is the right choice if there are few dozens of the instances of the minority class. Yu et al. suggested cost-sensitive learning methods attributed to modifying the training datasets to handle the CIP [15]. The strategy proposed in the study by Yu et al. [15] is to assign weights to the training data instance during

resampling by the cost decision matrix. Cheng et al. [16] employed cost-sensitive learning by changing the learning process to build a classifier using imbalanced data [16]. This study used the strategy to modify the support vector machine's objective function by weight updating procedure. Wang and Yao [17] proposed cost-sensitive boosting to handle the CIP.

2.3 Studies addressing CIP

The CIP has been dealt with in literature in real-life engineering problems. In software engineering predictive modeling, CIP is addressed to improve defect prediction models and models predicting change prone classes. Wang and Yao [17] inspected resampling techniques, threshold moving, and ensembles to handle the CIP. The study was carried out on datasets extracted from ten real-life software engineering projects. Khoshgoftaar et al. [18] investigated different feature selection scenarios and data resampling methods to improve software defect prediction models. Their work confirmed the effectiveness of feature selection based on resampled data to predict faulty classes on imbalanced datasets. Tan et al. [19] analyzed resample approaches and updatable classification to develop better-quality defect-prediction models. Pelayo and Dick [20] studied the effect of oversampling techniques for software defect prediction. Sun et al. [21] given a coding-based scheme to handle CIP in fault prediction. In this study, a coding-based scheme is applied to extract binary class data out of a multiclass dataset to train the fault prediction models using ML techniques.

Siers and Islam [22] provided cost-sensitive classification as a potential solution to CIP by proposing CSforest as a cost-sensitive classification technique, which is an ensemble of decision trees. The study also proposed a cost-sensitive voting scheme to take advantage of a set of decision tree classifiers to minimize the classification cost. Laradji et al. [23] investigated the optimistic effects of feature selection with ensemble learning to develop effective models for fault prediction with imbalanced data. Zheng [24] proposed a cost-sensitive boosting algorithm to boost the neural network performance for predicting defective classes in software. Choeikiwong and Vateekul [25] offered an unbiased version of the support vector machine (SVM) called R-SVM to predict defective modules in software with a high prediction rate. Malhotra and Khanna [26] studied resampling methods and meta-cost learning techniques to develop change prediction models using the imbalanced datasets. Thus, many studies in literature addressed the CIP in predicting change-prone and faulty classes. However, for SMP, the CIP has rarely been addressed in the literature. Thus, this study will be a pioneer in the prediction of software maintainability on imbalanced datasets. This empirical investigation employs three data resampling methods to balance the datasets and develops SMP models by applying ML techniques.

3 Design components

This section describes all the design components of this empirical study.

3.1 Independent and dependent variable

The development of the maintainability prediction model requires a dataset. The dataset consist of datapoints encompassing independent and dependent variables. The independent variables capture and characterize the various aspect of software systems such as complexity, size, coupling, and cohesion.

3.1.1 Independent variables

The following OO metrics (independent variables) are used in this paper. These metrics quantify various characteristics a software system.

Chidamber and Kemerer (CK) metrics: CK metrics used as independent variables in this paper include, Number of Children (NOC), Weighted Methods of a Class (WMC), Coupling between Objects (CBO), Response for a Class (RFC), Depth of Inheritance Tree (DIT), Lack of Cohesion in Methods (LCOM) For developing software quality models, these metrics are used in various studies [27,28].

Quality Model for Object-Oriented Design metric suite (QMOOD): The study uses few metrics from this metric suite, namely Number of Public Methods (NPM) Cohesion Among Methods of a class (CAM), Measure of Aggression (MOA), Method of Functional Abstraction (MFA), Data Access Metric (DAM) for developing SMP models.

Miscellaneous metrics: Few other metrics have also been used in this study as independent variables in addition to the above OO metrics. These metrics are Efferent Coupling (Ce), Afferent Coupling (Ca) given by [29] and Average Method Complexity (AMC), Coupling between methods of a class (CBM), Lines of code (SLOC), and LCOM3 given by [30]. These metrics capture the structural complexity of the software classes. In addition to the above metrics, Inheritance Coupling (IC) is also included as the independent variable. The URL for the detailed description of the software metrics used in the study is to see Gromit.iiar.pwr.wroc.pl/p_inf/ckjm/metric.html website.

3.1.2 Dependent variable

In this study, the dependent variable is "maintainability effort" which is a binary variable having two values: high and low. The high values of maintainability effort of a class indicate that the corresponding class will require high effort with respect to code modification during the maintenance and the class in which this value is low, will require less effort. Consequently, high maintainability effort classes necessitate more time and resources during maintenance. In this study, we aim to build competent SMP models that predict high maintainability effort classes early in the software development process. This way, these classes can be appropriately tested during testing and can be given proper resources in the maintenance period.

3.2 Subject selection

We have extracted the datasets for empirical validation from ten Apache software systems, namely Apache Io, Ivy, Betwixt, Java caching system (Jcs), Lang, Ode, and Log4j, and Byte Code Engineering Library (Bcel), Accumulo and Click. Their brief explanation is as follows:

Apache Bcel gives a very convenient way to produce, manipulate, and investigate the Java class files. The objects

represent the classes and provide all necessary symbolic information about the corresponding classes. *Apache Betwixt* provides users a convenient way of converting the beans into Extensible Markup Language (XML) and generating digester rules. A rule is fired when a pattern is matched. *Apache Io* is a library of utilities that assist in input and output functionality. It includes utility classes, filter classes, file monitoring, and comparator classes. *Apache Ivy* is an application package to manage project dependencies. It includes tracking, recording, and resolving project dependencies and is characterized by its flexibility. Written in Java language, *Apache Jcs* is a Java caching system which tends to manage the cache data to speed up applications. It acts like a front-tier cache that is configured to maintain consistency across multiple web servers. *Apache Lang* provides various utilities like basic numerical methods, string manipulation, and serialization of different system properties. *Apache Log4j* is a fast and reliable logging package with three primary components: logger, appender, and layout. The logger captures the logging information. Appender is accountable for publishing the logging information, whereas the layout component formats the information in different formats. *Apache Ode* software package designed to talk to the web services by sending and receiving messages, handle error recovery, and data manipulation. Apache *Accumulo* is a scalable and distributed data store. It is based on Google's Bigtable design and one of the most popular NoSQL database management system. Unlike traditional relational database management system, it provides distributed key-value storage that does not operate on rows. *Apache Click* is a web application framework that uses an event-based programming model to process servlet requests. It provides a rich client style programming model to commercial Java developers.

As the selected projects are open-source in nature, the study can easily be replicated by other researchers, reducing the threat to external validity.

3.3 Performance metrics

This section illustrates various performance metrics that we have used in this study to evaluate the SMP models. When the dataset is imbalanced, appropriate performance metrics are required to examine the predictive capability of the developed models. The researchers have suggested a few robust metrics for CIP. He and Garcia [31] proposed that G-Mean (geometric mean of negative and positive classes) metric should be used for the imbalanced. Gao et al. [32] suggested that Balance is a robust performance metric if the dataset is imbalanced [32]. Lessmann et al. [33] supported the use of Receiver Operator Characteristics (ROC) analysis as in useful performance metric for imbalanced data. In this analysis, Area under Curve (AUC) is advocated as a valuable metric to assess the prediction models. This study evaluates the SMP models with the help of G-Mean and AUC. In addition to G-Mean and AUC, we also used traditional performance metrics, namely accuracy, specificity, sensitivity, and precision, to evaluate the SMP models. All of these metrics can be calculated from the confusion matrix. The confusion matrix for a binary classification problem is shown in Table 1. In this study, positives are considered as the classes requiring high maintainability effort, whereas the contrary is the classes requiring low maintainability effort. A brief definition of the performance metrics, along with the formulas, is given below.

Table 1 Confusion matrix

	Predicted positive	Predicted negative
Actual positive	True positive (TPos)	False Negative (FNeg)
Actual negative	False positive (FPos)	True negative (TNeg)

Accuracy: The percentage of correctly predicted negative and positive data points.

$$\text{Accuracy} = \frac{TPos + TNeg}{TPos + TNeg + FPos + FNeg}, \tag{1}$$

Sensitivity: The percentage of correctly predicted positive data points.

$$\text{Sensitivity} = \frac{TPos}{TPos + FNeg}, \tag{2}$$

Specificity: The percentage of correctly predicted negative data points.

$$\text{Specificity} = \frac{TNeg}{TNeg + FPos}, \tag{3}$$

Precision: The percentage of correctly predicted positive out of total positive data points.

$$\text{Precision} = \frac{TPos}{TPos + FPos}, \tag{4}$$

G-Mean: The geometric mean of Specificity and Sensitivity.

$$\text{G-Mean} = \sqrt{\text{Sensitivity} \times \text{Specificity}}. \tag{5}$$

Area under receiver operator characteristics curve (AUC): receiver operator characteristics curve is plotted between 1-specificity and sensitivity taking 1-specificity and sensitivity as x and y coordinate, respectively.

3.4 Statistical tests

To strengthen the conclusions and validate the hypothesis formed in the study, we use statistical tests. The Friedman test is conducted for AUC and G-Mean to ascertain whether the performance of prediction models differs significantly after applying a particular data resampling method. Depending upon the performance metric used for evaluation of a specific method of resampling, the Friedman test assigns a mean rank as per the performance of that resampling technique. After knowing from the Friedman test that the data resampling techniques used in the study have significantly improved the performance of SMP models, we extended our statistical investigation beyond with the Wilcoxon test.

4 Experimental framework

This section is dedicated to data collection procedure, data preprocessing, and model validation criteria.

4.1 Empirical data collection

The study uses ten open-source project datasets briefly described in Section 3.2. The datasets are extracted with the assistance of the Defect Collection and Reporting (DCRS) tool [34] that is developed in house at Delhi Technological

University. The requirement for data extraction is using DCRS tool is that the repository must use version control using git. Since the projects selected for this study are various application packages of Apache that use git version control, so datasets corresponding to all ten application packages are collected with the DCRS tool. We fed two consecutive versions of the software projects as an input to the DCRS tool and retrieved the change logs in the form of log records for the classes that were the same or common between the two versions. The log record keeps the data regarding the change descriptions, list of modified files, and list of lines of code changed. The CKJM tool that is an open-source software metric extractor is also embedded in the DCRS tool to determine the OO metrics corresponding to the subject under selection in this study. The data points are generated for the classes that were common between the two versions. The OO metrics and change statistics corresponding to a specific class form a data point. The change statistics describe the lines of code added, deleted, modified, and total changes (CHANGE), which is the sum of all three. The data points also include an "ALTER" variable containing two values "Yes" and "No." The "Yes" value signifies that change in lines of code has happened in that class, and "No" means that there is no change corresponding to that class. The data points in which "ALTER" contained "Yes" are retained in the dataset, and the remaining data points were discarded. When all of the Apache application packages are analyzed with the DCRS tool, we got data points containing the values of OO metrics and CHANGE. The details of the projects after data collection are shown in Table 2. For each software project, we show the number of common classes (#C), the number of common classes changed (#CC), and the change percentage (C%) in Table 2.

4.2 Data pre-processing

Data preprocessing is carried out in two steps: outlier analysis and data discretization. If a data point significantly deviates from the data set's remaining data points, such a data point is called an outlier. It is important to identify such data points and filters them out so that a competent predictive model can be developed. In this direction, outlier analysis is carried out in the study using the interquartile filter available in WEKA (Waikato Environment for Knowledge Analysis) tool. The details of the datasets after outlier analysis has given in Table 3. These outliers are removed from the datasets before developing the prediction models. The dependent variable maintainability effort is formed after the discretization of continuous variable CHANGE into two bins.

The data points in the two bins are assigned high and low values for maintainability effort. The class in which the value of maintainability is high indicates that the class will demand more effort in terms of changing lines of source code during the maintenance phase, and the class in which this value is low, will require less effort. The data points for which maintainability effort is high are regarded as minority class instances. The number of majority (#Maj.) and the number of minority class (#Min.) instances in all datasets used in the study are reported in Table 4. The imbalance ratio (IR) in the datasets used in the study ranged from 4.52 to 23.80, indicating the imbalanced nature of these datasets. Here, the imbalance ratio is the number of majority class instances divided by the number of minority class instances. The number of data points in each dataset after outlier analysis has given in Table 3.

4.3 The data balancing with data resampling methods

The next phase involves using data resampling methods to balance the data to provide uniform class distribution of minority and majority class instances for efficiently training the prediction model. In this study, we apply three resampling methods to balance the data. We assess the performance SMP models after each of the resampling techniques.

4.4 Development and evaluation of the SMP models

The last phase is developing and evaluating the SMP models. We apply seven ML techniques on ten datasets under consideration in the study to develop models. Further, we train the SMP models using the TFCV method. With TFCV, the

Table 3 Number of outliers is each dataset

Dataset	No. of outliers
Betwixt	34
Ivy	58
Bcel	26
Io	24
Log4j	32
Jcs	22
Lang	47
Ode	115
Accumulo	90
Click	31

Table 2 Details of software projects

Software	Version	#C	#CC	C%
Betwixt	0.6 – 0.7	290	279	96.20
Ivy	1.4.1 – 2.2.0	619	429	69.30
Bcel	5.0 – 5.1	363	360	99.17
Io	2.0.1 – 2.2	274	272	99.27
Log4j	1.2.14 – 1.2.15	491	437	89.00
Jcs	1.2.6.5 – 1.2.7.9	333	219	65.76
Lang	2.4 – 2.6	434	267	61.52
Ode	1.3.1 – 1.3.2	1060	1004	94.71
Accumulo	1.5.0-1.5.1	1431	828	57.86
Click	2.2-2.3	402	263	65.42

Table 4 Data discretization results

Dataset	#Min.	#Maj.	IR
Betwixt	27	218	8.07
Ivy	28	343	12.25
Bcel	18	316	17.55
Io	10	238	23.80
Log4j	42	363	8.64
Jcs	27	170	6.29
Lang	27	193	7.14
Ode	57	832	14.59
Accumulo	120	618	5.15
Click	42	190	4.52

dataset available for training the model is split up into ten equal-sized subsets. Randomly, nine subsets of the training data are taken for training, whereas the leftover subset is utilized for testing the model.

The technique iterates ten times, repeating the same procedure. In this way, each partition is used for validating the model once. Once the SMP models are trained for the designated task, the next step involves their evaluation. To do so, we use performance measures AUC and G-Mean, as discussed in Section 3.3. Apart from AUC and G-Mean, we also evaluate the SMP models using accuracy, precision, sensitivity, and specificity as well. Furthermost, to compare the performance of different techniques to handle the imbalanced data, we perform statistical tests. The application of statistical test strengthens the conclusions of the study, thereby enhancing the conclusion validity. The experimental framework is diagrammatically depicted in Fig. 1.

5 Research methodology

This section describes the data resampling methods and ML techniques used to develop the maintainability prediction models.

5.1 Data resampling methods

The data resampling techniques widely acknowledged in the literature are used in this study to deal with CIP. The data resampling is the data level solution to CIP that tries to balance the training datasets by oversampling and undersampling. The oversampling techniques reduce the strength of the majority class by removing the data points, whereas Undersampling tends to increase the strength of the minority class by generating new data points. These techniques are quite adaptable because they can be used independently of the selected learning technique. Before training the SMP models using ML techniques, we preprocessed the imbalanced datasets by data resampling. Since this study is the first to address the CIP in software maintainability prediction, we leverage the data level techniques already established for software engineering predictive modeling and analyze the applicability of these techniques for SMP. The datasets under consideration in the study are imbalanced, i.e., classes requiring high maintainability effort are very less in number compare to classes requiring low maintainability effort. So, it is difficult to develop effective models to predict the class requiring high maintainability effort accurately. Therefore, in this study, we apply three established data resampling methods, namely: SM, RR, and SS, to get rid of CIP. SM is an oversampling technique that works by increasing the instances of the minority class to achieve a uniform distribution of instances of the two classes [35]. SS is an undersampling method which makes a balance in the data set by randomly confiscating the majority class data points while the RR can serve both. We briefly describe the resampling methods as follows:

SS: This technique adjusts the distribution of the two classes by removing the data points of the majority class [19]. Removing the data points of the majority class helps in balancing the ratio of two classes. For the SS technique in this

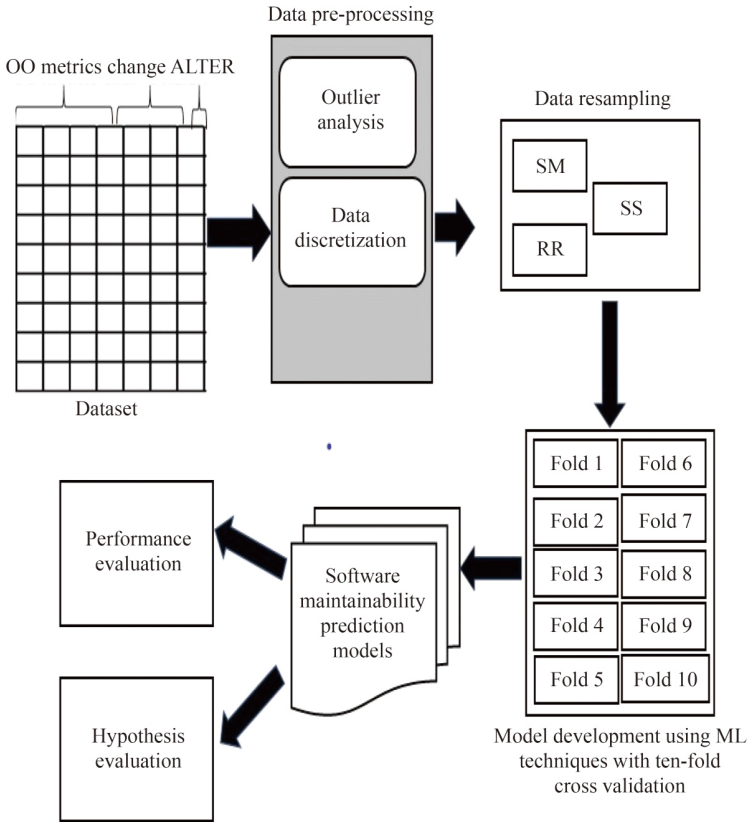


Fig. 1 Experimental framework

study, we assess different distribution ratios ranging from 1:1 to 10:1 and select the best ratio.

SM: This is an oversampling method for balancing the imbalanced datasets. The technique works by creating artificial minority class data points and including them in the dataset. The artificial data points along the line segment joining a randomly selected minority class sample with all or any of its *k*-nearer neighbors are created and added in the training dataset to oversample the minority class. We set the value of $k = 5$ for SM. Then, the neighbors of minority class instances are randomly picked to achieve a required amount of oversampling. For example, if 300 percent oversampling is needed, three neighbors out of five of a particular minority class instance are randomly picked, and one artificial sample is generated corresponding to each neighbor in the direction of that neighbor. The synthetic instances are created using the following: The initial phase includes computing the difference between the picked minority class data point and its closest neighbor. In the next step, the computed difference by a random number somewhere in the range of 0 and 1. The resultant multiplication is added to the minority class data point under consideration. We examine five oversampling percentages (100%, 200%, 300%, 400%, and 500%) for SM in the study and select the best rate.

RR: This technique is based on bootstrap methodology. The number of subsamples of the minority class is created by with replacement strategy to increase the distribution of the minority class. A different sample is created every time and added to the training dataset [19]. Since this method is very random, we perform the ten iterations of the method and report the average results.

5.2 ML techniques

In this section, we briefly describe the ML techniques used in the study to develop maintainability prediction models. The ML techniques used in the study are Naïve Bayes (NB), Logit Boost (LB), Multilayer Preceptor (MLP), Random Forest (RF), J48 Decision Tree, Bagging (BG), and Adaptive Boosting (AB). For simulation, the study uses the WEKA tool (https://www.cs.waikato.ac.nz/ml/weka/) with default parameter settings.

NB is a probability-based classifier and based on Bayes Theorem. It relies on the assumption that predictor variables are independent, given the class, and no hidden input feature impact the prediction process. These assumptions make it a competent technique in predictive modeling. Being a probabilistic classifier, it highly compensates for class imbalance [36]. In default parameter settings of the WEKA tool, NB kernel estimator and supervised discretization are kept false.

MLP is a feed-forward back propagation artificial neural network consisting input layer, one or more hidden layers, and an output layer of neurons. In this technique, the learning process is carried out in two phases, namely the forward and backward pass. In the forward pass, input features are presented to the input layer. Each neuron in the hidden layer then transforms the input features to the weighted summation, followed by an activation function. The output layer finally produces the output after receiving the input from the hidden layers. The difference between the actual and predicted outputs is termed as an error. If the predicted output matches the desired output, the network's performance is considered good otherwise in the backward pass, the error is propagated back, and network readjusts the weights [37]. In the WEKA tool, the default parameter setting of MLP includes momentum equal to 0.2, the validation threshold equal to 20, and the learning rate equal to 0.3.

RF is a Meta classifier that fits some decision trees. Various subsamples are created out of the input features with replacement, and the decision tree is fitted to each subsample. The RF classifier's output is the average output produced by individual decision tree classifiers [38]. In the default parameter setting for the RF, the number of decision trees is equal to 100.

AB is based on the boosting strategy. Boosting is an ensemble technique that creates a robust classifier by using several weak classifiers. AB technique improves the performance of weak classifiers, and it works as follows: The weak classifiers are learned by using the weighted samples of the training data. The misclassification rate of an individual trained model is calculated, and the training instance weights are modified in such a way that the correctly classified instances get lower weights while the incorrectly classified instance gets the higher weight. In this way, the focus of the algorithm is more on the instances which are challenging to learn [39]. For model development using the WEKA tool, the default parameter settings for AB are: Decision Stump as the weak classifier with ten numbers of iterations and a weighted threshold of 100.

BG is based on a bootstrap approach. The technique works by creating many subsamples of the training dataset. The resampling is done with replacement to avoid overfitting. The classifiers are developed on individual subsamples of the training instances. The output of the bagged classier is the average of the outputs of the different classifiers [40]. For model development using the WEKA tool, the default parameter settings for BG are: Bag size = 100, the base classifier is REPTree, and the number of iterations = 10.

LB is a boosting technique that boosts the sequence of classifiers iteratively to produce a competent and robust model. In each iteration, the training samples are reweighed to improve the classification error [41]. In the WEKA tool default parameter setting for this technique, the base classifier is decision stump, likelihood threshold = −1.79, and the number of iterations = 10.

J48 is a decision tree that practices entropy for attribute selection measure [36].

6 Results and analysis

This section depicts the answers to research questions designed in the paper.

6.1 RQ1) How are the predictions of SMP models built with datasets that are imbalanced?

Table 5 shows the performance of SMP models developed on imbalanced data sets with the application of ML techniques. The performance of the models is examined using perfor-

mance metrics: sensitivity, precision, specificity, accuracy, AUC, and G-Mean. In analyzing Table 5, it is found that SMP developed using imbalanced data has given good accuracy. In the majority of the cases, accuracy values are higher than 90%. However, using the accuracy metric used to evaluate the performance of classification models has been criticized when the underlying dataset is imbalanced [31,32]. Also, it is quite evident from Table 5 that sensitivity values are quite low and varies from 0 to 40% in the majority of the cases.

Similarly, the precision values are also quite low, ranging from 0 to 40% in most of the cases. Lower sensitivity and precision values are because prediction models have not learned the minority class instances appropriately. Another observation from Table 5 is that the specificity values are above 95% in the majority of the cases. Thus, all these trends, i.e., low sensitivity, precision, and high specificity values of developed prediction models are because datasets are imbalanced in nature and prediction models have not learned the minority class instances (maintainability effort = high) appropriately.

6.2 RQ2) Do the data resampling methods results in improved prediction models?

To answer this RQ, we apply data resampling methods on ten imbalanced datasets and then develop SMP models with the application of seven ML techniques. The performance of developed maintainability prediction models is evaluated using performance measures accuracy, sensitivity, precision, specificity, G-Mean, and AUC. For SM, five oversampling percentages stated in Section 5 were examined for each dataset, and the best performance is reported. Similarity for SS technique, we examined ten different distribution ratios for individual data set and reported the best performance metric values corresponding to that dataset. Figures 2−5 shows the average values of accuracy, sensitivity, specificity, and precision of SMP models developed after data resampling. As shown in Fig. 2, the average accuracy of SMP models developed after data resampling using the RR technique is higher than 80% for seven out of ten datasets. For SM technique, average accuracy values reported being greater than 80% for eight datasets out of ten datasets, whereas SMP models developed after SS technique gave accuracy greater than 70% in seven out of ten datasets. After data resampling, the results of SMP models in terms of sensitivity are also encouraging for RR and SM techniques. For RR technique,

Table 5 Results of SMP models without data resampling

Dataset	Accuracy							Sensitivity						
	NB	MLP	AB	BG	LB	RF	J48	NB	MLP	AB	BG	LB	RF	J48
Bcel	85.03	94.01	94.01	93.41	95.51	94.31	93.41	16.67	22.22	11.11	5.56	22.22	11.11	22.22
Betwixt	83.67	87.35	88.98	87.76	90.72	86.53	85.71	18.52	22.22	0.00	0.00	3.70	7.41	3.70
Io	94.76	96.37	95.97	95.16	98.74	94.38	95.16	40.00	30.00	30.00	0.00	20.00	27.27	0.00
Ivy	90.79	92.72	91.91	91.96	92.72	91.11	91.91	30.77	17.86	3.75	3.57	14.29	10.71	14.29
Jcs	83.76	91.88	87.31	89.34	86.80	88.32	88.32	77.78	62.96	48.15	44.44	40.74	48.15	48.15
Lang	90.00	89.55	90.91	90.91	91.36	90.45	90.91	44.44	51.85	66.67	55.56	55.56	48.15	51.85
Log4j	86.67	90.12	87.16	88.15	90.86	89.63	87.90	47.62	33.33	26.19	9.52	26.19	35.71	23.81
Ode	90.10	93.81	94.15	93.93	94.15	92.46	92.91	36.84	22.81	8.77	10.53	12.28	10.53	22.81
Accumulo	80.89	81.84	83.76	83.06	84.01	81.57	82.93	24.17	20.00	0.00	9.17	7.50	15.83	9.17
Click	77.59	81.03	79.74	81.47	78.88	79.74	79.31	4.76	21.43	0.00	0.00	2.38	19.05	2.38
Dataset	Precision							Specificity						
	NB	MLP	AB	BG	LB	RF	J48	NB	MLP	AB	BG	LB	RF	J48
Bcel	07.89	40.00	33.33	16.67	80.00	40.00	33.33	88.92	85.03	98.73	98.42	99.68	99.05	97.47
Betwixt	21.74	37.50	0.00	0.00	14.29	20.00	10.00	91.74	95.41	100	98.62	98.11	96.33	95.87
Io	36.36	60.00	50.00	0.00	40.00	33.33	0.00	97.06	99.16	98.74	99.16	98.74	97.48	99.16
Ivy	33.33	55.56	25.00	25.00	57.14	27.27	40.00	95.34	98.83	99.13	99.13	99.13	97.67	98.25
Jcs	44.68	73.91	54.17	66.67	52.38	59.09	59.09	84.71	96.47	93.53	96.47	94.12	94.71	94.71
Lang	63.16	58.33	62.07	65.22	68.18	65.00	66.67	96.37	94.82	94.30	95.85	96.37	96.37	96.37
Log4j	38.46	53.85	34.38	28.57	64.71	50.00	37.04	91.18	96.69	94.21	97.25	98.35	95.87	95.32
Ode	28.77	54.17	100	66.67	77.78	27.27	40.63	93.75	98.68	100	99.64	99.76	98.08	97.72
Accumulo	36.71	38.71	0.00	40.74	56.25	35.19	39.29	91.91	93.85	100	97.41	98.87	94.34	97.25
Click	14.29	45.00	0.00	0.00	11.11	35.19	0.00	93.68	94.21	97.37	99.47	95.79	93.16	96.32
Dataset	AUC							G-Mean						
	NB	MLP	AB	BG	LB	RF	J48	NB	MLP	AB	BG	LB	RF	J48
Bcel	0.57	0.73	0.78	0.75	0.79	0.79	0.62	38.50	46.69	52.37	23.38	47.07	33.17	46.54
Betwixt	0.69	0.65	0.72	0.69	0.68	0.7	0.54	52.11	46.05	0.00	0.00	19.06	26.71	18.84
Io	0.87	0.87	0.87	0.85	0.85	0.79	0.57	62.31	54.54	54.43	0.00	44.44	51.56	0.00
Ivy	0.67	0.66	0.78	0.72	0.74	0.77	0.51	54.16	42.01	18.42	18.82	37.63	32.35	37.46
Jcs	0.86	0.91	0.89	0.87	0.88	0.71	0.71	81.17	77.94	67.11	65.48	61.92	67.53	67.53
Lang	0.83	0.83	0.80	0.86	0.83	0.68	0.83	65.45	70.12	79.29	72.97	73.17	68.12	70.69
Log4j	0.84	0.77	0.78	0.83	0.79	0.77	0.7	65.89	56.77	49.67	30.43	50.75	58.51	47.64
Ode	0.78	0.76	0.77	0.80	0.79	0.75	0.63	58.77	47.44	29.62	32.39	35.00	32.13	47.21
Accumulo	0.62	0.69	0.67	0.69	0.72	0.75	0.54	47.13	43.32	0.00	29.88	27.23	38.65	29.86
Click	0.57	0.64	0.66	0.66	0.67	0.67	0.47	21.12	44.93	0.00	0.00	15.10	42.12	15.14

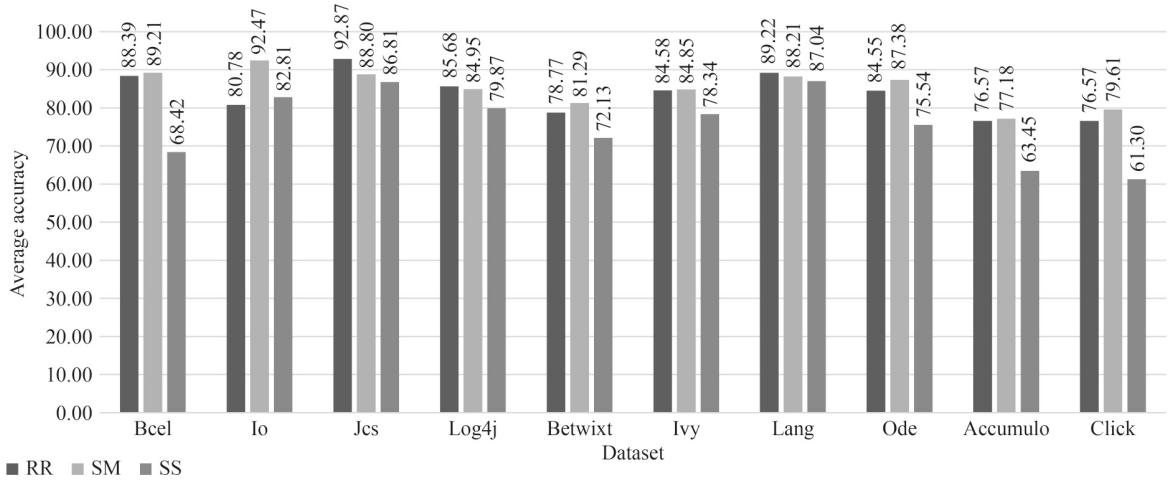


Fig. 2 Average accuracy after data resampling

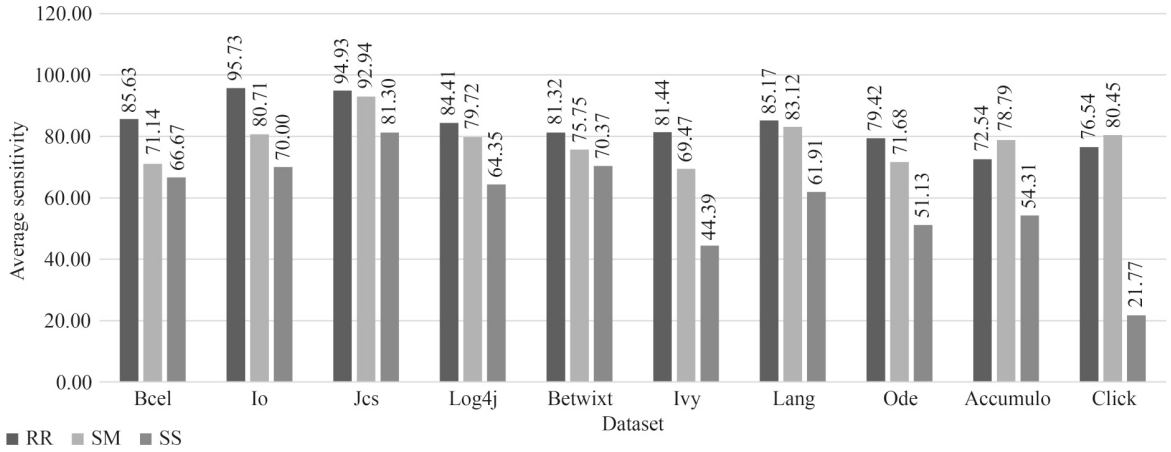


Fig. 3 Average sensitivity after data resampling

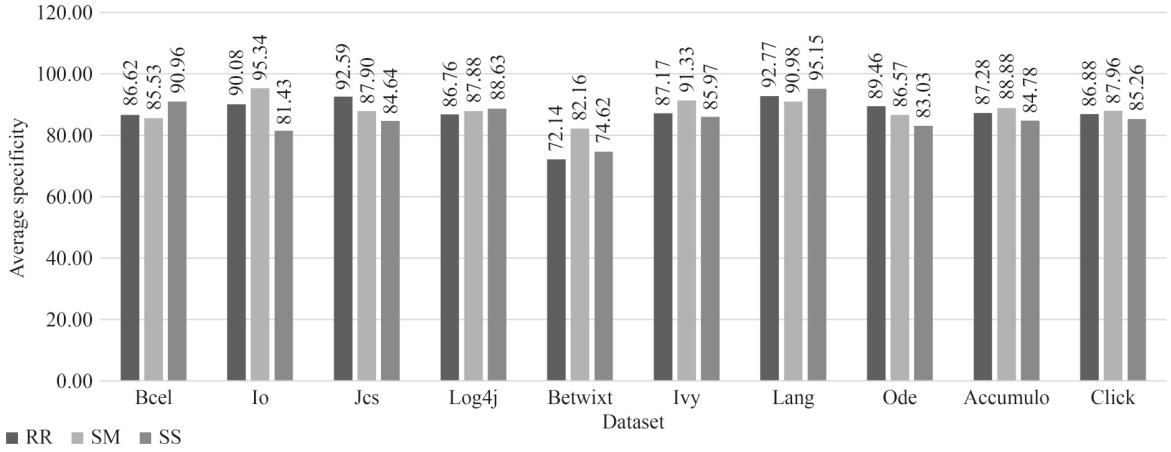


Fig. 4 Average specificity after data resampling

average sensitivity (Fig. 3) reported being greater than 80% for seven datasets and for SM technique, average sensitivity reported being greater than 70% respectively for nine out of ten datasets used in the study, whereas for SS, average sensitivity reported being greater than 60% in six datasets. Similarly, on analyzing the performance of SMP models developed after data resampling, average specificity (Fig. 4) was reported to be higher than 80% for nine datasets for the RR technique. For SM technique, average specificity was computed to be higher than 80% for all datasets, whereas for the SS technique, the average specificity was higher than 80% for nine out of ten datasets. The performance of SMP models in terms of precision is also found to be very encouraging for RR and SM techniques (Fig. 5). For RR the average precision

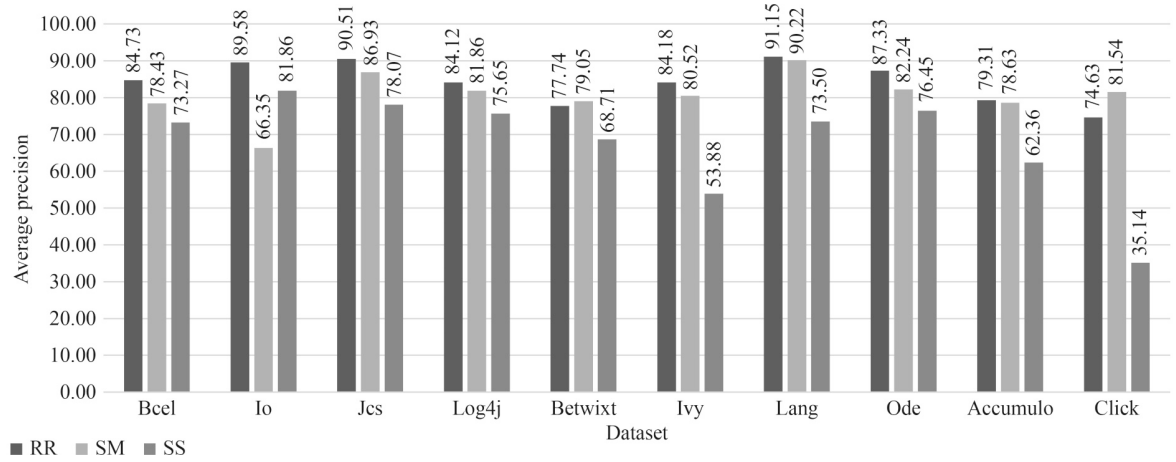


Fig. 5 Average Precision after data resampling

reported being greater than 80% in seven datasets and for SM technique, average precision was greater than 70% in nine datasets. Whereas, for SS, SMP models gave average precision greater than 60% for eight out of ten datasets. For performance measure, G-Mean (Table 6), and AUC (Table 7), SMP models gave an excellent performance when data resampling methods are employed. The majority of the cases' G-Mean after data resampling ranged from 70% to 98.02%. The G-Mean after data resampling with RR techniques are greater than 80 in 71.4% of the cases whereas, for SM technique, G-Mean values are higher than 80 in 62.85% of the cases. The AUC values are higher than 0.80 for 90% and 91%

Table 6 G-Mean results after data resampling

Dataset	Classifier	RR	SM	SS	Dataset	Classifier	RR	SM	SS
	NB	61.85	64.84	70.28		NB	52.12	52.81	53.29
	MLP	86.00	78.76	68.94		MLP	83.31	82.10	70.37
	AB	80.45	78.05	80.51		AB	37.56	77.68	71.62
Bcel	BG	93.19	82.59	76.80	Betwixt	BG	84.97	84.73	73.98
	LB	89.30	75.52	89.24		LB	75.98	84.46	72.20
	RF	95.27	75.36	86.78		RF	92.73	76.71	85.15
	J48	92.90	88.01	69.92		J48	89.01	85.09	74.07
	NB	81.58	87.16	57.45		NB	49.80	62.89	52.49
	MLP	92.89	87.74	79.37		MLP	90.19	80.26	67.61
	AB	94.10	87.28	79.37		AB	79.97	71.98	63.39
Io	BG	93.70	88.26	83.67	Ivy	BG	90.41	87.44	64.14
	LB	95.71	85.83	73.48		LB	79.09	76.11	63.39
	RF	96.11	87.84	73.48		RF	95.82	90.22	41.41
	J48	94.91	89.93	77.14		J48	93.97	85.67	69.85
	NB	89.57	84.36	85.65		NB	69.50	72.16	74.07
	MLP	96.51	91.88	70.27		MLP	91.35	89.15	72.98
	AB	87.40	90.91	88.52		AB	88.54	87.50	77.48
Jcs	BG	94.02	90.98	78.26	Lang	BG	92.33	89.18	67.94
	LB	96.10	88.56	86.86		LB	89.77	93.83	82.32
	RF	98.02	92.79	83.15		RF	96.21	83.72	87.41
	J48	94.48	93.10	83.15		J48	90.98	90.42	73.17
	NB	78.35	75.53	72.17		NB	62.61	67.63	68.22
	MLP	91.24	85.55	69.01		MLP	86.48	79.37	51.88
	AB	73.43	79.36	76.19		AB	72.19	79.47	71.45
Log4j	BG	90.50	85.07	72.84	Ode	BG	92.33	86.21	75.35
	LB	79.51	82.57	80.99		LB	82.71	54.86	35.34
	RF	94.08	90.35	76.64		RF	96.05	91.18	70.67
	J48	90.35	87.27	78.51		J48	91.66	87.64	70.87
	NB	45.16	61.36	50.73		NB	42.78	65.24	31.72
	MLP	74.98	73.43	59.41		MLP	81.39	79.22	47.43
	AB	45.16	71.55	74.98		AB	68.24	78.90	38.98
Accumulo	BG	89.86	85.12	61.18	Click	BG	84.61	81.30	34.50
	LB	74.63	74.76	67.76		LB	75.00	85.78	46.38
	RF	94.65	88.96	64.95		RF	87.35	84.82	46.29
	J48	89.37	82.42	59.11		J48	83.99	79.66	44.22

Table 7 AUC results after data resampling

Dataset	Classifier	RR	SM	SS	Dataset	Classifier	RR	SM	SS
Bcel	NB	0.78	0.78	0.76	Betwixt	NB	0.77	0.82	0.69
	MLP	0.93	0.9	0.77		MLP	0.89	0.89	0.75
	AB	0.97	0.91	0.77		AB	0.79	0.87	0.79
	BG	0.97	0.95	0.83		BG	0.93	0.93	0.78
	LB	0.94	0.93	0.93		LB	0.85	0.92	0.8
	RF	0.99	0.96	0.84		RF	0.97	0.95	0.82
	J48	0.96	0.9	0.72		J48	0.93	0.88	0.74
Io	NB	0.95	0.97	0.78	Ivy	NB	0.72	0.75	0.7
	MLP	0.94	0.97	0.91		MLP	0.93	0.9	0.74
	AB	0.96	0.97	0.88		AB	0.89	0.86	0.8
	BG	0.98	0.95	0.80		BG	0.97	0.95	0.78
	LB	0.98	0.96	0.89		LB	0.9	0.89	0.81
	RF	0.99	0.96	0.89		RF	0.99	0.96	0.78
	J48	0.95	0.89	0.76		J48	0.95	0.85	0.77
Jcs	NB	0.91	0.89	0.89	Lang	NB	0.83	0.87	0.83
	MLP	0.97	0.95	0.92		MLP	0.95	0.92	0.82
	AB	0.95	0.94	0.89		AB	0.95	0.93	0.78
	BG	0.97	0.97	0.86		BG	0.96	0.93	0.83
	LB	0.98	0.94	0.91		LB	0.95	0.95	0.82
	RF	0.98	0.98	0.89		RF	0.96	0.95	0.82
	J48	0.96	0.94	0.89		J48	0.93	0.9	0.69
Log4j	NB	0.88	0.86	0.82	Ode	NB	0.84	0.81	0.83
	MLP	0.96	0.93	0.84		MLP	0.91	0.89	0.75
	AB	0.86	0.89	0.83		AB	0.87	0.9	0.81
	BG	0.95	0.94	0.82		BG	0.97	0.95	0.81
	LB	0.89	0.91	0.82		LB	0.88	0.93	0.81
	RF	0.97	0.96	0.84		RF	0.99	0.97	0.9
	J48	0.93	0.92	0.79		J48	0.94	0.89	0.74
Accumulo	NB	0.74	0.70	0.713	Click	NB	0.62	0.68	0.51
	MLP	0.83	0.80	0.64		MLP	0.87	0.86	0.6
	AB	0.83	0.78	0.65		AB	0.73	0.86	0.52
	BG	0.96	0.92	0.68		BG	0.91	0.9	0.54
	LB	0.811	0.78	0.70		LB	0.81	0.93	0.59
	RF	0.99	0.94	0.70		RF	0.95	0.92	0.95
	J48	0.94	0.87	0.64		J48	0.9	0.86	0.57

for data resampling with RR and SM, respectively. For SS, in 51.40% of the cases, AUC values reported being higher than 0.80. These positive improvements indicate that maintainability prediction models can correctly predict minority class after data resampling.

6.3 RQ3) Among the data resampling methods employed in this study, which method performs best to handle the imbalanced data?

To answer this RQ, we perform the Friedman test on all ten data sets used in the study concerning stable performance measures G-Mean and AUC for three data sampling methods along with the situation when no resampling is employed. The following hypothesis is tested.

H_0: *Null hypothesis*: The G-Mean and AUC results of the SMP models developed with ML techniques do not show any significant difference before and after data resampling.

H_a: *Alternate hypothesis*: The G-Mean and AUC results of the SMP models developed with ML techniques have significant difference before and after data resampling.

We evaluate the above hypothesis at $\alpha = 0.05$ with three degrees of freedom. We apply the Freidman test by extracting the performance measure G-Mean and AUC for individual datasets for all ML techniques used in the study. The higher the rank achieved by a particular resampling method, the better the resampling method. Tables 8 and 9 shows the results of the Friedman test with respect to G-Mean and AUC, respectively.

Corresponding to each technique, the mean rank achieved by resampling methods along with no resampling situation is stated in Tables 8 and 9.

According to Table 8, the RR method outperforms on all ten datasets used in the study concerning G-Mean, and the results are significant (p-value < 0.05) in all ten datasets used in the study. Similarly, as shown in Table 9, RR technique outperforms on the nine of the ten datasets used in the study with respect to AUC, and the results are significant (p-value < 0.05). Thus, the null hypothesis is rejected, and it is concluded that the use of data resampling methods significantly improves SMP models' performance compared to when no resampling is employed. The RR method is emerged as an excellent resampling technique as per the ranking obtained by the Friedman test. Also, it is quite evident from Table 8, and Table 9 worst rank was obtained when no resampling is used. Subsequently, we say that use of data resampling improves the predictive capability of SMP models. As shown in Tables 8

Table 8 Friedman test results for G-Mean

Data set (GM)	Rank 1	Rank 2	Rank 3	Rank 4	p-value
Apache Bcel	RR (3.57)	SM (2.57)	SS (2.86)	No Resample (1.00)	0.002
Apache Betwixt	RR (3.29)	SM (3.14)	SS (2.57)	No Resample (1.00)	0.003
Apache IO	RR (3.86)	SM (3.14)	SS (1.86)	No Resample (1.14)	0.000
Apache Ivy	RR (3.57)	SM (3.14)	SS (2.00)	No Resample (1.29)	0.003
Apache Jcs	RR (3.71)	SM (3.00)	SS (2.14)	No Resample (1.14)	0.001
Apache Lang	RR (3.57)	SM (3.00)	SS (2.14)	No Resample (1.29)	0.006
Apache Log4j	RR (3.43)	SM (3.29)	SS (2.29)	No Resample (1.00)	0.001
Apache Ode	RR (3.57)	SM (3.14)	SS (2.29)	No Resample (1.00)	0.001
Apache Accumulo	RR (3.14)	SM (2.29)	SS (2.43)	No Resample (1.14)	0.007
Apache Click	RR (3.57)	SM (3.43)	SS (2.00)	No Resample (1.00)	0.000

Table 9 Friedman test results AUC

Data set (AUC)	Rank 1	Rank 2	Rank 3	Rank 4	p-value
Apache Bcel	RR (3.93)	SM (3.00)	SS (1.93)	No Resample (1.14)	0.000
Apache Betwixt	SM (3.57)	RR (3.36)	SS (2.00)	No Resample (1.07)	0.000
Apache IO	RR (3.57)	SM (3.43)	SS (1.71)	No Resample (1.29)	0.001
Apache Ivy	RR (3.86)	SM (3.14)	SS (2.00)	No Resample (1.00)	0.000
Apache Jcs	RR (3.86)	SM (3.07)	SS (1.86)	No Resample (1.21)	0.001
Apache Lang	RR (3.64)	SM (3.21)	No Resample (1.86)	SS (1.29)	0.001
Apache Log4j	RR (3.71)	SM (3.29)	SS (1.71)	No Resample (1.29)	0.001
Apache Ode	RR (3.71)	SM (3.14)	SS (2.00)	No Resample (1.14)	0.001
Apache Accumulo	RR (4.00)	SM (2.86)	No Resample (1.71)	SS (1.43)	0.001
Apache Click	RR (3.50)	SM (3.29)	No Resample (1.71)	SS (1.50)	0.003

and 9, the RR method outperforms than the other resampling methods; we further carry the Wilcoxon test at the significance level; $\alpha = 0.05$ to discover if the RR method is statistically superior to other data resampling techniques. The test is carried out concerning the performance measures AUC and G-Mean of all ML techniques for all datasets.

H_{01}: The RR technique has the same performance terms of AUC and G-Mean as that of SM and SS techniques.

H_{a1}: The RR technique has not the same performance terms of AUC and G-Mean as that of SM and SS techniques.

The p-values after the pairwise comparison of the RR technique with SM and SS for the Wilcoxon signed-rank test are depicted in Table 10. The p-values shown in Table 10 are less than or equal to 0.05 for all four examined pairs. We reject the null hypothesis H_{01}, which states SMP models developed by balancing datasets with the RR technique have the same performance in terms of AUC and G-Mean as that of models developed by balancing datasets with SM and SS. We conclude that SMP models developed after resampling with RR are statistically better in their performance compared to SM and SS methods to handle CIP to develop competent SMP models.

7 Validity threats

We express different potential validity threats of the study in this section.

Conclusion validity: Improper analysis of the results of the empirical study create a threat to conclusion validity, which is also termed as the threat to statistical validity [42]. This threat does not exist in this study as we conducted a proper statistical analysis to support our conclusions. Also, SMP models developed in this study are validated using TFCV, which even lower down the validation bias and conclusion validity threat.

Internal validity: In context to internal validity, it is to be noted that the balanced datasets obtained after applying the data resampling methods are not an accurate representation of the original dataset. When this balanced dataset is used to learn the SMP models, we may not have generalized it for testing. This may lead to validation bias and a potential threat to internal validity. However, the study uses the robust performance measure AUC and G-Mean to analyze predictions of SMP models. These performance measures make a suitable balance between the true positive and true negative predictions, reducing the validation bias.

Construct validity: Several software quality predictive modeling studies have successfully validated the predictors that we have used in this investigation. Thus, the construct validity threat about the predictors used in this study does not exist. The software maintainability is the dependent variable in this study. This variable is formed after discretization of total lines of code changed. The lines of code changes is calculated by examining the change logs by the DCRS tool, whose accuracy assured. Therefore, with respect to response variable there is no threat to construct validity.

External validity: The study uses ten datasets, which represent the ten Apache open-source software written in the Java language. The results may vary datasets obtained from the software system of other programming languages. Hence, a threat to external validity exits in this study. However, the study default parameter setting to develop maintainability

Table 10 Wilcoxon test results

Pairwise resampling method	G-Mean	AUC
RR vs. SM	Sig+(0.005)	Sig+(0.006)
RR vs. SS	Sig+(0.00)	Sig+(0.00)

prediction models, which reduces the threats to generalizability.

8 Discussion

The maintainability effort is estimated is SLOC revised in the maintenance phase. The SLOC revised include lines of code added, deleted and modified. The classes having high maintainability effort involve in a greater number of revisions in the source code in comparison to those of low maintainability effort classes during the maintenance. The study is intended in developing competent prediction models to forecast high maintainability effort classes by treating imbalanced datasets. In this study, we have only regarded changes that occur due to defects in a class. This means a class with a higher number of defects or more complex defects may require more SLOC changes as compared to another class in which defects can be resolved with the low number of SLOC changes. Thus, a high number of SLOC changes indicate the less maintainable class or class requiring high maintainability effort and vice versa. The higher number of defects in a class indicates poor design, which is an indication of low maintainability or high maintainability effort. We have gathered historical data pertaining to changes based on defects and developed models that can be used to predict the classes that are prone to more changes in the early phases of software development. The prediction of high maintainability effort classes with the aid of SMP models aid in reducing the fault-proneness of such classes. The high maintainability effort classes can be documented well to enhance their understandability when carrying out the future maintenance activities. Also, before releasing the software, the developers can refactor classes requiring high maintainability effort to improve their maintainability i.e., the classes predicted to be of requiring high maintainability can be refactored by the developers at earlier stages rather than the maintenance programmers to refactor them late during the maintenance phase as developers are more knowledgeable about the interior structure of the code of the classes than the external maintenance programmers. Many studies have shown that code cloning also increases maintenance effort [43−45] due to increases in the code growth, enhancement dead code and chances of introduction of bugs. Therefore, the classes predicted to be requiring high maintainability effort with the developed models can be refactored to eliminate the cloning.

9 Conclusions and future work

SMP models help developers predicting the maintainability of software components at the early stages of software development so that the limited available resources of the organization can be allocated optimally. This study is a successful effort in the direction of developing efficient SMP Models. Ten imbalanced open-source datasets and seven ML techniques are used in this study to develop SMP models. The study uses three data resampling methods to deal with CIP. We used TFCV to validate the SMP models. The results of the study are statistically examined using statistical tests. The major conclusions are summarized as follows:

- The use of data resampling methods, namely RR, SM, and SS, significantly improve the predictions of the SMP models.
- The RR method outperformed other resampling methods, namely SM and SS, for handling imbalanced datasets when the results of the developed models are analyzed for G-Mean and AUC.
- By the Friedman test, the SM method is emerged as the second-best method to handle imbalanced datasets because SM has attained the second-highest rank on all datasets concerning the performance measures G-Mean and AUC.

Thus, this study supports the RR method to develop efficient SMP models when the dataset in hand is imbalanced. The developed maintainability prediction models can effectively be used by the software practitioners to predict maintainability even if the dataset is imbalanced. In the future, we will investigate few algorithm level methods and meta cost learning approaches in dealing with CIP.

References

1. Kaur A, Kaur K. Statistical comparison of modelling methods for software maintainability prediction. International Journal of Software Engineering and Knowledge Engineering, 2013, 23(6): 743–774
2. Kumar L, Rath S K. Hybrid functional link artificial neural network approach for predicting maintainability of object-oriented software. Journal of Systems and Software, 2016, 121: 170–190
3. Malhotra R, Lata K. An exploratory study for predicting maintenance effort using hybridized techniques. In: Proceedings of the 10th Innovations in Software Engineering Conference. 2017, 26–33
4. Kumar L, Naik D K, Rath S K. Validating the effectiveness of object-oriented metrics for predicting maintainability. Procedia Computer Science, 2015, 57: 798–806
5. Elish M O, Elish K O. Application of TreeNet in predicting object-oriented software maintainability: a comparative study. In: Proceeding of the 13th European Conference on Software Maintenance and Reengineering. 2009, 69–78
6. Chug A, Malhotra R. Benchmarking framework for maintainability prediction of open source software using object oriented metrics. International Journal of Innovative Computing, Information and Control, 2016, 12(2): 615–634
7. Thwin M M T, Quah T S. Application of neural networks for software quality prediction using object-oriented metrics. Journal of Systems and Software, 2005, 76(2): 147–156
8. Van Koten C, Gray A R. An application of Bayesian network for predicting object-oriented software maintainability. Information and Software Technology, 2006, 48(1): 59–67
9. Aggarwal K K, Singh Y, Kaur A, Malhotra R. Application of artificial neural network for predicting maintainability using object-oriented metrics. International Journal of Computer and Information Engineering, 2008, 2(10): 3552–3556
10. Zhang W, Huang L, Ng V, Ge J. SMPLearner: learning to predict software maintainability. Automated Software Engineering, 2015, 22(1): 111–141
11. Kumar L, Rath S K. Software maintainability prediction using hybrid neural network and fuzzy logic approach with parallel computing concept. International Journal of System Assurance Engineering and Management, 2017, 8(S2): 1487–1502
12. Zhou L. Performance of corporate bankruptcy prediction models on imbalanced dataset: the effect of sampling methods. Knowledge-Based Systems, 2013, 41: 16–25

13. Napierala K, Stefanowski J. Types of minority class examples and their influence on learning classifiers from imbalanced data. Journal of Intelligent Information Systems, 2016, 46(3): 563–597
14. Loyola-González O, Martínez-Trinidad J F, Carrasco-Ochoa J A, García-Borroto M. Study of the impact of resampling methods for contrast pattern based classifiers in imbalanced databases. Neurocomputing, 2016, 175: 935–947
15. Yu H, Mu C, Sun C, Yang W, Yang X, Zuo X. Support vector machine-based optimized decision threshold adjustment strategy for classifying imbalanced data. Knowledge-Based Systems, 2015, 76: 67–78
16. Cheng F, Zhang J, Wen C. Cost-sensitive large margin distribution machine for classification of imbalanced data. Pattern Recognition Letters, 2016, 80: 107–112
17. Wang S, Yao X. Using class imbalance learning for software defect prediction. IEEE Transactions on Reliability, 2013, 62(2): 434–443
18. Khoshgoftaar T M, Gao K, Seliya N. Attribute selection and imbalanced data: problems in software defect prediction. In: Proceedings of the 22nd IEEE International Conference on Tools with Artificial Intelligence. 2010, 137–144
19. Tan M, Tan L, Dara S, Mayeux C. Online defect prediction for imbalanced data. In: Proceedings of the 37th IEEE International Conference on Software Engineering. 2015, 99–108
20. Pelayo L, Dick S. Applying novel resampling strategies to software defect prediction. In: Proceedings of 2007 Annual Meeting of the North American Fuzzy Information Processing Society. 2007, 69–72
21. Sun Z, Song Q, Zhu X. Using coding-based ensemble learning to improve software defect prediction. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2012, 42(6): 1806–1817
22. Siers M J, Islam M Z. Software defect prediction using a cost sensitive decision forest and voting, and a potential solution to the class imbalance problem. Information Systems, 2015, 51: 62–71
23. Laradji I H, Alshayeb M, Ghouti L. Software defect prediction using ensemble learning on selected features. Information and Software Technology, 2015, 58: 388–402
24. Zheng J. Cost-sensitive boosting neural networks for software defect prediction. Expert Systems with Applications, 2010, 37(6): 4537–4543
25. Choeikiwong T, Vateekul P. Software defect prediction in imbalanced data sets using unbiased support vector machine. In: Kim K J. ed. Information Science and Applications. Berlin, Heidelberg: Springer, 2015, 923−931
26. Malhotra R, Khanna M. An empirical study for software change prediction using imbalanced data. Empirical Software Engineering, 2017, 22(6): 2806–2851
27. Giger E, Pinzger M, Gall H C. Can we predict types of code changes? An empirical analysis. In: Proceedings of the 9th IEEE Working Conference on Mining Software Repositories. 2012, 217–226
28. Singh Y, Kaur A, Malhotra R. Empirical validation of object-oriented metrics for predicting fault proneness models. Software Quality Journal, 2010, 18(1): 3–35
29. Martin R C. Agile Software Development: Principles, Patterns, and Practices. Upper Saddle River: Prentice Hall, 2002
30. Henderson-Sellers B. Object-Oriented Metrics: Measures of Complexity. Upper Saddle River: Prentice Hall, 1995
31. He H, Garcia E A. Learning from imbalanced data. IEEE Transactions on Knowledge and Data Engineering, 2009, 21(9): 1263–1284
32. Gao K, Khoshgoftaar T M, Napolitano A. Combining feature subset selection and data sampling for coping with highly imbalanced software data. In: Proceedings of the 27th International Conference on Software Engineering and Knowledge Engineering. 2015, 439–444
33. Lessmann S, Baesens B, Mues C, Pietsch S. Benchmarking classification models for software defect prediction: a proposed framework and novel findings. IEEE Transactions on Software Engineering, 2008, 34(4): 485–496
34. Malhotra R, Pritam N, Nagpal K, Upmanyu P. Defect collection and reporting system for Git based open source software. In: Proceedings of 2014 International Conference on Data Mining and Intelligent Computing. 2014, 1–7
35. Chawla N V, Bowyer K W, Hall L O, Kegelmeyer W P. SMOTE: synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 2002, 16: 321–357
36. Han J, Pei J, Kamber M. Data Mining: Concepts and Techniques. 3rd ed. New York: Elsevier, 2011
37. Haykin S. Neural Networks: A Comprehensive Foundation. Upper Saddle River: Prentice Hall, 1994
38. Breiman L. Random forests. Machine Learning, 2001, 45(1): 5–32
39. Witten I H, Frank E. Data Mining: Practical Machine Learning Tools and Techniques. 2nd ed. San Francisco: Morgan Kaufmann, 2005
40. Breiman L. Bagging predictors. Machine Learning, 1996, 24(2): 123–140
41. Friedman J, Hastie T, Tibshirani R. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). The Annals of Statistics, 2000, 28(2): 337–407
42. Ma Y, Luo G, Zeng X, Chen A. Transfer learning for cross-company software defect prediction. Information and Software Technology, 2012, 54(3): 248–256
43. Nguyen H A, Nguyen T T, Pham N H, Al-Kofahi J, Nguyen T N. Clone management for evolving software. IEEE Transactions on Software Engineering, 2012, 38(5): 1008–1026
44. Duala-Ekoko E, Robillard M P. Tracking code clones in evolving software. In: Proceedings of the 29th International Conference on Software Engineering. 2007, 158–167
45. De Wit M, Zaidman A, Van Deursen A. Managing code clones using dynamic change tracking and resolution. In: Proceedings of 2009 IEEE International Conference on Software Maintenance. 2009, 169–178

Ruchika Malhotra is Associate Head and Associate Professor in the Discipline of Software Engineering, Department of Computer Science and Engineering, Delhi Technological University (formerly Delhi College of Engineering), India. She has been awarded the prestigious UGC Raman Postdoctoral Fellowship by the Indian government for pursuing postdoctoral research from Department of Computer and Information Science, Indiana University-Purdue University, USA. She was an assistant professor at the University School of Information Technology, Guru Gobind Singh Indraprastha University, India. She has received IBM Faculty Award 2013. Her h-index is 30 as reported by Google Scholar. She can be contacted by email at ruchikamalhotra@dtu.ac.in.

Kusum Lata is currently working as Assistant Professor in University School of Management and Entrepreneurship, Delhi Technological University, India. She completed her master's degree in Computer Technology and Applications in 2010 from the Delhi College of Engineering, University of Delhi, India. She has received her doctoral degree in Computer Science and Engineering from Delhi Technological University, India. Her research interests are in software quality improvement, applications of machine learning techniques in maintainability prediction, and validation of software metrics. She can be contacted by email at kusumlata@dtu.ac.in.

molecules

MDPI

Review

Insights into Structural Modifications of Valproic Acid and Their Pharmacological Profile

Manish Kumar Mishra [1,2], Samiksha Kukal [1,3], Priyanka Rani Paul [1,3], Shivangi Bora [1,2], Anju Singh [4,5], Shrikant Kukreti [4], Luciano Saso [6], Karthikeyan Muthusamy [7], Yasha Hasija [2] and Ritushree Kukreti [1,3,*]

1 Genomics and Molecular Medicine Unit, Institute of Genomics and Integrative Biology (IGIB), Council of Scientific and Industrial Research (CSIR), Mall Road, Delhi 110007, India; manishkrmishra18@gmail.com (M.K.M.); samiksha@igib.in (S.K.); priyankapaul008@gmail.com (P.R.P.); shivee30@gmail.com (S.B.)
2 Department of Biotechnology, Delhi Technological University, Shahbad Daulatpur, Main Bawana Road, Delhi 110042, India; yashahasija06@gmail.com
3 Academy of Scientific and Innovative Research (AcSIR), Ghaziabad 201002, India
4 Nucleic Acids Research Lab, Department of Chemistry, University of Delhi (North Campus), Delhi 110007, India; anju11278@gmail.com (A.S.); skukreti@chemistry.du.ac.in (S.K.)
5 Department of Chemistry, Ramjas College, University of Delhi (North Campus), Delhi 110007, India
6 Department of Physiology and Pharmacology "Vittorio Erspamer", Sapienza University of Rome, P. le Aldo Moro 5, 00185 Rome, Italy; luciano.saso@uniroma1.it
7 Department of Bioinformatics, Alagappa University, Karaikudi 630004, Tamil Nadu, India; mkbioinformatics@gmail.com
* Correspondence: ritus@igib.in or ritus@igib.res.in; Tel.: +91-11-27662202; Fax: +91-11-27667471

Abstract: Valproic acid (VPA) is a well-established anticonvulsant drug discovered serendipitously and marketed for the treatment of epilepsy, migraine, bipolar disorder and neuropathic pain. Apart from this, VPA has potential therapeutic applications in other central nervous system (CNS) disorders and in various cancer types. Since the discovery of its anticonvulsant activity, substantial efforts have been made to develop structural analogues and derivatives in an attempt to increase potency and decrease adverse side effects, the most significant being teratogenicity and hepatotoxicity. Most of these compounds have shown reduced toxicity with improved potency. The simple structure of VPA offers a great advantage to its modification. This review briefly discusses the pharmacology and molecular targets of VPA. The article then elaborates on the structural modifications in VPA including amide-derivatives, acid and cyclic analogues, urea derivatives and pro-drugs, and compares their pharmacological profile with that of the parent molecule. The current challenges for the clinical use of these derivatives are also discussed. The review is expected to provide necessary knowledgebase for the further development of VPA-derived compounds.

Keywords: valproic acid; anticonvulsant; analogues; derivatives; teratogenicity; hepatotoxicity; development

check for updates

Citation: Mishra, M.K.; Kukal, S.; Paul, P.R.; Bora, S.; Singh, A.; Kukreti, S.; Saso, L.; Muthusamy, K.; Hasija, Y.; Kukreti, R. Insights into Structural Modifications of Valproic Acid and Their Pharmacological Profile. *Molecules* **2022**, *27*, 104. https://doi.org/10.3390/molecules27010104

Academic Editor: Enrico Millo

Received: 27 November 2021
Accepted: 18 December 2021
Published: 24 December 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

Valproic acid (VPA, 2-propyl-pentanoic acid), is a fatty acid derivative of valeric acid naturally found in *Valeriana officinalis* [1,2]. Synthesized in 1881 by Beverly S Burton [3], VPA continued to be used as an organic solvent till 1963 before its role as an anticonvulsant was discovered serendipitously by Meunier et al. [4]. Following its first marketing as an antiepileptic drug (AED) in France, VPA represents one of the most efficient and widely prescribed AED world-wide [1]. With its recognized role as an AED, VPA has drastically evolved as a molecule targeting a spectrum of pathologies such as bipolar disorder, migraines and neuropathic pain [5–7]. VPA has shown promise as a neuroprotective agent in Alzheimer's disease and multiple sclerosis [8–10]. Additionally, overlapping pathways and interactions have led to its recognition as a compound of interest in the clinical trials for a wide variety of cancers [11–14]. However, the potential adverse effects associated with

VPA use are hepatotoxicity and teratogenicity, which interfere with its therapeutic use [15]. Substantial experimentation and research have gone into elimination of these adverse effects and to improve its efficacy by introducing structural changes in the VPA skeletal structure. These changes ranged from branching and alpha fluorination to production of chiral compounds, from its amide derivatives to cyclic derivatives and conjugation products. This article briefly discusses on the structural elements of valproic acid, its pharmacokinetics and pharmacodynamic profile. The focus of the review is to provide a comprehensive account of all the derivatives and analogues of VPA and ascertain their success as possible replacement of the parent compound as better, safe and tolerable drugs. Any compound second generation to VPA shall address the current challenges of potency and teratogenicity along with hepatotoxicity of the metabolites. This knowledge of the existing modifications will serve a platform to design better successors of VPA.

2. Structure and Pharmacology of Valproic Acid

2.1. VPA Structural Elements

The simple chemical structure of VPA makes it unique in the arsenal of drugs that are used for same treatments as VPA. It is an eight-carbon molecule with the backbone of pentanoic acid and a propyl group attached to the second carbon. It is a weak organic acid with pKa value of 4.8, that causes 99.8% ionization of the molecule at physiological pH [16]. Therefore, despite of good lipophilicity (logP = 2.75) it requires carrier protein for its cellular uptake [16]. Keane et al., reported lack of anticonvulsant activity of non-branched analogues of VPA in vivo [17]. On the contrary, branched analogues were found to significantly increase the brain concentration of gamma-aminobutyric acid (GABA), showing anticonvulsant activity against induced seizures in animal models [17–19]. These studies indicated the essentiality of the branched structure of VPA for anticonvulsant activity. It has also been proposed that C=O group is the main functional group controlling the anticonvulsant behavior by providing the initial electrostatic interaction between VPA and the target protein [20]. Other than the molecule's therapeutic effect, structure-teratogenicity relationship studies performed in mouse strains implicated that the tetrahedral structure, the hydrogen attached to C-2 and aliphatic side chain branching [21] may be responsible for the undesirable teratogenic effect associated with its use [22]. Lloyd et al., also suggested the role of carboxylic acid moiety in causing neural tube defect. It has been reported that the carboxylic group of VPA that exists as an anion at physiological pH interacts with organic anion transporters (OATs) present on the foetal facing/basolateral plasma membrane of the placenta [23]. This interaction of VPA with OATs leads to VPA accumulation in foetus causing embryotoxicity [23]. Next substantial side-effect, hepatotoxicity is primarily caused by metabolites of VPA (4-ene-VPA and 2,4-diene-VPA) [24–27].

2.2. VPA Pharmacokinetics

Valproic acid is available for treatment in various forms such as oral solutions, tablets, enteric-coated capsule and intravenous solution, with bioavailability up to 100% and absorption half-life ranging from 30 min to 4 h depending upon the type of formulation used [28,29]. Among the present preparations, the oral formulations are widely used with enteric coated capsule showing highest bio-availability compared to other oral formulations [28]. Food intake influences the absorption of VPA, with absorption being slowest when taken after three hours of a meal as compared to when taken with meals or during fasting conditions [30].

Established metabolic routes of VPA are glucuronidation and beta oxidation, which account for 50% and 40% of VPA metabolism, respectively [31,32]. Studies have also explored the involvement of cytochrome P450 (CYP450)-mediated oxidation, although its contribution remains low (~10% VPA metabolism) [33,34]. A major metabolite found in urine samples is a conjugate product of VPA and glucuronic acid (15% to 40% of total VPA administered in humans) [35,36]. The metabolites produced via oxidation pathways are known to cause hepatotoxicity. Direct esterification of VPA or esterification of one

of the major metabolites, 4-ene-VPA, forms respective acetyl-CoA esters [37,38]. These esters are highly reactive and they form thiol conjugates with mitochondrial glutathione, which result in the depletion of mitochondrial glutathione. The resulting reduction in the glutathione pool causes interference with mitochondrial function, which leads to hepatotoxicity [37,39,40].

The plasma level of a free fraction of VPA therapeutic concentration is 6%; the rest is bound to protein mainly albumin. Due to its high affinity for the protein, VPA shows a low clearance rate of 6–8 mL/h/kg [41–43].

2.3. VPA Pharmacodynamics

Valproic acid is primarily prescribed for the treatment of epilepsy, bipolar disorder and migraines. The anticonvulsant property of VPA is due to its influence on brain concentrations of GABA. It indirectly inhibits GABA degradation by hindering the activity of enzymes, such as GABA transaminase (GABA-T) and succinate semialdehyde dehydrogenase (SSADH) that are involved in GABA decay [44]. Furthermore, VPA acts on GABA-receptor to potentiate postsynaptic inhibitory responses [29].

Other than affecting inhibitory synapses, VPA has been demonstrated to reduce the neuron surface expression and synaptic localization of glutamate-receptor subunits, GluR1/2, thus preventing excitatory effect [29]. It also obstructs various voltage gated ion channels such as Na^+ channels, voltage-gated K^+ channels, and voltage-gated Ca^{2+} channels [45]. Such an obstruction relaxes the high frequency neuronal firing [46,47]. The downregulation of sodium voltage-gated channel alpha subunit 3 (SCN3a) expression by reducing the gene promoter methylation in mouse neural crest-derived cells is also suggested as an epigenetic pathway in the anticonvulsant role of VPA [48]. More recent finding of VPA pharmacodynamics is its role as histone deacetylase (HDAC) inhibitor [12,13]. leading to increase in gene expression of apoptotic proteins and hence its activity in cancer treatment [14]. Apart from these mechanisms, evidence demonstrates that VPA affects other signaling systems to exert its antimanic, neuroprotective, anti-inflammatory and antioxidant effect, which may lead to extension of its use in other non-epileptic neurological disorders; however, discussing these diverse roles is beyond the scope of this article. Figure 1 summarizes major molecular targets of VPA for various therapeutic areas.

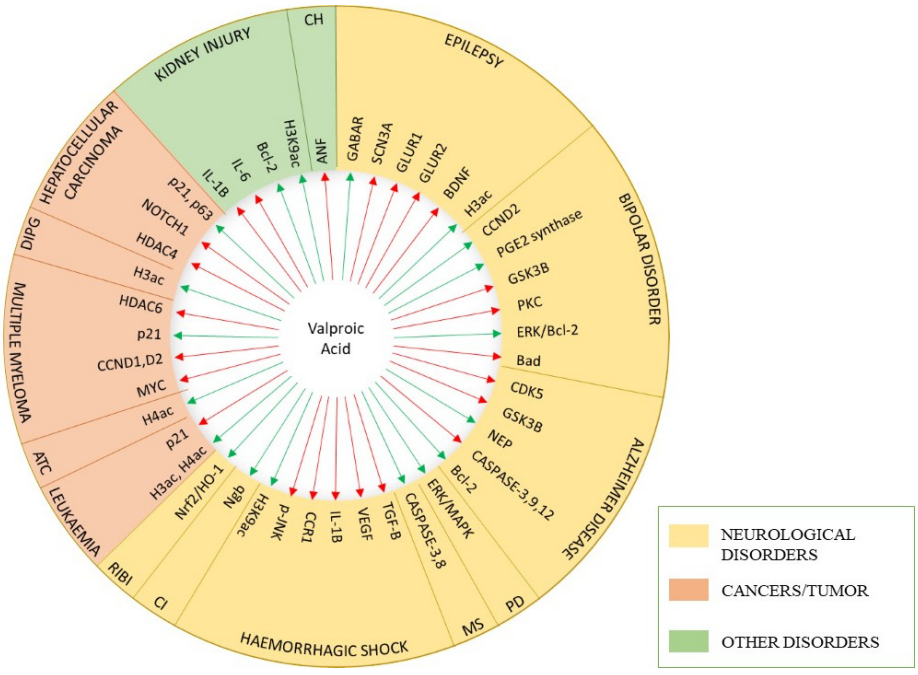


Figure 1. Molecular targets of valproic acid. Different targets regulated by valproic acid are represented by arrows for neurological disorders (epilepsy, bipolar disorder, Alzheimer's disease, Parkinson's

disease (PD), encephalomyelitis), haemorrhagic shock, ischemia, injury (radiation-induced brain injury (RIBI), kidney injury), cardiac hypertrophy (CH) and cancers (leukaemia, anaplastic thyroid carcinoma (ATC), multiple myeloma, diffuse intrinsic pontine glioma (DIPG), hepatocellular carcinoma). The green arrows indicate upregulation of expression and red arrows indicate inhibition/downregulation of target gene.

3. Structural Modification of Valproic Acid: Derivatives and Analogues

In attempts to minimize the adverse effects associated with VPA use and to improve its potency, various analogues and derivatives have been synthesized and studied in comparison to the parent molecule. As shown in Figure 2, primary modifications include altering the carboxylic acid group (replacing it with the amide); altering the main chain (replacing with cyclic entities) and/or changing the branching pattern by introducing unsaturated bonds. This section elaborates the literature reported under each VPA derivative describing the structural changes to the parent molecule, comparison of pharmacological properties in terms of potency, activity and toxicity profile and the clinical applicability of the compound. A brief summary of the above is shown in Table 1.

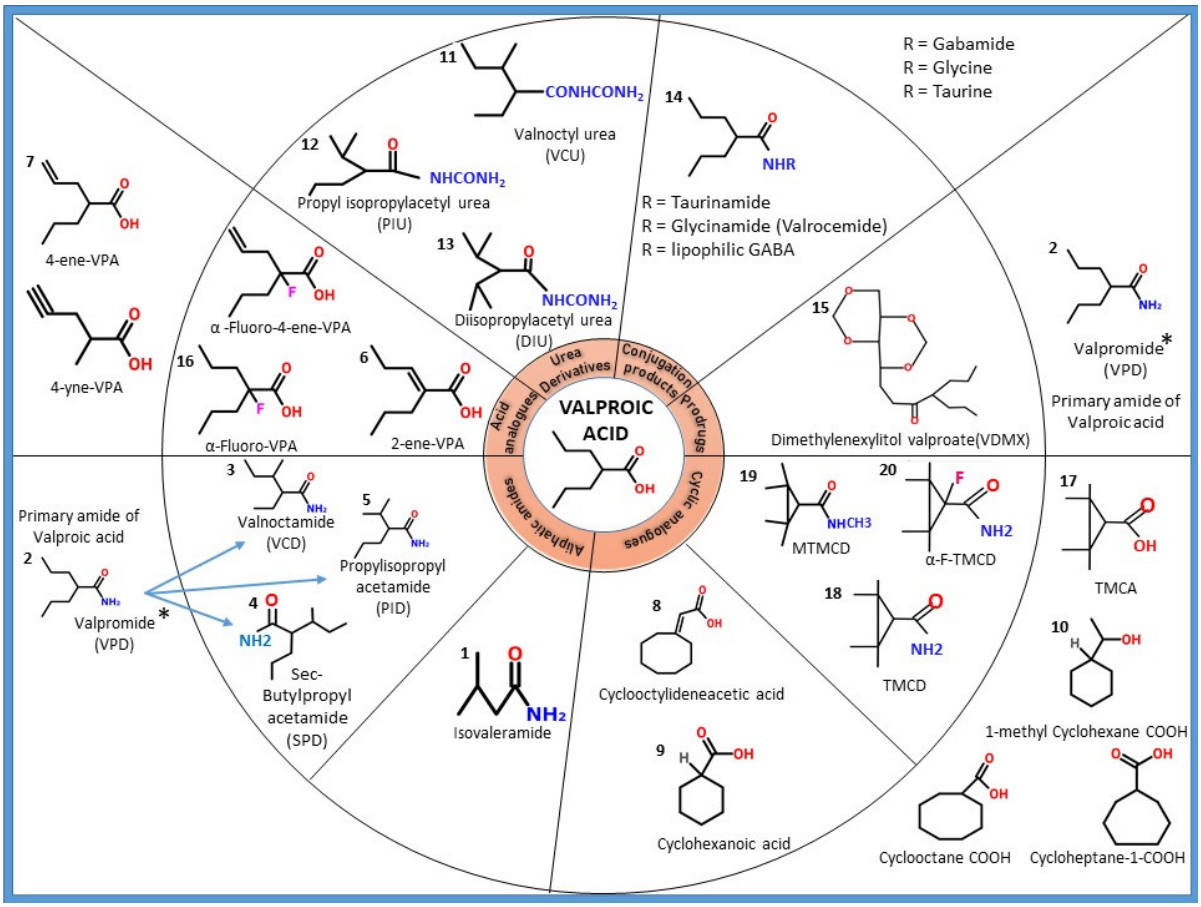


Figure 2. Schematic representation of valproic acid derivatives and analogues. These compounds are characterised into aliphatic amides, acid analogues, conjugation products, prodrugs and cyclic analogues. The aliphatic amides are further characterised into chiral and achiral compounds, and acid analogues are further characterised into unsaturated and alpha fluorinated compounds. The compounds outside the circle are those with some limitations ceasing their further study. The compounds included in the circle are those that are under active investigation. * Valpromide is mentioned twice since along with being a primary amide of valpromide; it also behaves as a prodrug.

Table 1. Valproic acid analogues and derivatives, classification and pharmacological activity.

S. No	Candidate Compound	Parent Compound	Classification	Chemical Formula	Molecular Weight (g/mol)	Anticonvulsant Property	Clinical Trial Phase	Teratogenicity	Hepatotoxicity	Reference
1	Isovaleramide	Isovaleric acid	Aliphatic amide	C5H11NO	101.149	Reduced	Phase I, [a] Phase II	Reduced	N.D.	[49]
2	Valpromide	Valproic acid	Aliphatic amide	C8H17NO	143.23	Increased	NA	* No change in humans	* No change in humans	[50]
3	Valnoctamide	Valpromide	Aliphatic amide	C8H17NO	143.23	Increased	[b] Phase III completed	Reduced	Reduced	[50]
4	sec-Butyl-propyl acetamide	Valpromide	Aliphatic amide	C9H19NO	157.25	Increased	Phase II	Reduced	Reduced	[51]
5	Propyl isopropyl acetamide	Valpromide	Aliphatic amide	C8H17NO	143.23	Increased	NA	Reduced	N.D.	[52]
6	2-ene-VPA	Valproic acid	Acid analogue	C8H14O2	142.2	Similar	[c] Phase I	Reduced	Reduced	[53]
7	4-ene-VPA	Valproic acid	Acid analogue	C8H14O2	142.2	Similar	NA	Increased	Increased	[27,54]
8	Cyclooctylideneacetic acid	Valproic acid	Cyclic analogue	C10H16O2	168.23	Increased	NA	N.D.	N.D.	[55]
9	Cyclohexane carboxylic acid	Valproic acid	Cyclic analogue	C7H12O2	128.17	Similar	NA	Reduced	N.D.	[55]
10	1-methyl cyclohexane carboxylic acid	Cyclohexane carboxylic acid	Cyclic analogue	C8H14O2	142.2	Increased	NA	N.D.	N.D.	[55]
11	Valnoctyl urea	Valnoctic acid	Urea Derivative	C9H18N2O2	186.25	Increased	NA	Reduced	N.D.	[56]
12	Propyl isopropylacetyl urea	Diisopropyl acetamide	Urea Derivative	C9H18N2O2	186.25	Increased	NA	Reduced	N.D.	[56]
13	Diisopropyl acetyl urea	Valproic acid	Urea Derivative	C9H18N2O2	186.25	Increased	NA	N.D.	N.D.	[56]
14	Valrocemide	Valproic acid	Conjugation product	C10H20N2O2	200.282	Increased	[d] Phase II	Absent	N.D.	[57,58]

Table 1. *Cont.*

S. No	Candidate Compound	Parent Compound	Classification	Chemical Formula	Molecular Weight (g/mol)	Anticonvulsant Property	Clinical Trial Phase	Teratogenicity	Hepatotoxicity	Reference
15	Dimethylenexylitol valproate	Valproic acid	Sugar ester	C15H26O6	302.36	Increased	NA	N.D.	N.D.	[59,60]
16	α-Floro-VPA	Valproic acid	Acid analogue	C8H15O2F	162.20	Reduced	NA	Reduced	Reduced	[61]
17	TMCA	Valproic acid	Cyclic analogue	C8H14O2	142.20	Reduced	NA	N.D.	N.D.	[62]
18	TMCD	TMCA	Cyclic analogue	C8H15NO	141	Increased	NA	Reduced	N.D.	[62,63]
19	MTMCD	TMCA	Cyclic analogue	C9H17NO	154	Increased	NA	Reduced	N.D.	[63,64]
20	α-Floro-TMCD	TMCA	Cyclic analogue	C8H14FNO	159.20	Increased	NA	Reduced	N.D.	[24,65]

Note—The change in anticonvulsant activity, teratogenicity and hepatotoxicity is mentioned with respect to valproic acid. * Since valpromide transforms into valproic acid in humans, there is no change in teratogenicity for valpromide; NA—not applicable, these compounds did not enter the clinical trials; N.D.—not determined; [a]: for acute migraine headache; [b]: for mania, schizoaffective disorder, manic type; [c]: drug tolerance study in health volunteers; [d]: for therapy-resistant patients with epilepsy.

3.1. Amide Derivatives of VPA

Amide derivatives of VPA have proven themselves to reduce the hepatotoxicity and teratogenicity of the parent compound. They have also depicted improved potency except valpromide, which systemically bio-transforms into VPA in human subjects [66].

3.1.1. Isovaleramide

Isovaleramide, a branched chain amide derivative of isovaleric acid (naturally occurring VPA analogue) with the molecular weight of 101.149 g/mol was developed by NPS Pharmaceuticals Inc., Salt Lake City, UT, USA under the alias of NPS 1776 [49]. Isovaleramide has a palatable taste and odour, unlike its parent compound isovaleric acid, which has an undesirable taste and odour [49]. Though, no mechanism of action for this compound has been reported, it is targeted for treatment of epilepsy and migraine attacks, and has shown broad spectrum of anticonvulsant activity in maximal electroshock stimulation (MES)-, pentylenetetrazole (PTZ)-induced seizures and kindling models along with chemo-convulsant and absence epilepsy models [49].

Islovaleramide exhibits linear pharmacokinetics in the dose range of 100–1600 mg [49]. The drug was found to be safe and well tolerated up to maximum dose of 2400 mg per day administered to phase 1 human volunteers with none of the adverse effects associated with VPA use [49]. Isovaleramide does not show any inhibitory effect on CYP450 enzymes, which may prevent any significant drug–drug interactions [49]. One limitation of the compound is its lower half-life of just 2.5 h, which leads to its rapid elimination. The reduced half-life of this potential drug can be targeted by making conjugation products or sustained release formulations [49]. Currently, the drug is undergoing phase II clinical trials for treatment of acute migraine headaches [67].

3.1.2. Valpromide (VPD)

Valpromide (VPD, dipropylacetamide, 2-propylvaleramide) is a 143.23 g/mol primary fatty acid amide of VPA, differing by the presence of amide group replacing the carboxylic acid functional group. This substitution makes it a more potent AED (2–5 times) as observed in in vivo models [68].

As opposed to the HDAC inhibitory effect of VPA, studies prove that VPD does not interfere with HDAC activity and thus has a highly diminished teratogenic effect in animals [12,69]. A study by France-Hélène Paradis and Barbara F. Hales [70] demonstrated that VPA causes downregulation of both SRY-Box Transcription Factor 9 (SOX9) and Runt-related transcription factor 2 (RUNX2)-signalling pathways involved in chondrogenesis and osteogenesis, leading to an overall teratogenic limb malformation in rat embryos. The downregulation of these pathways was likely the consequence of histone-4 hyperacetylation by VPA. VPD essentially failed at affecting histone 4 hyperacetylation or SOX9 expression, with only minimally altering the expression of RUNX2, thus lacking teratogenic potential. However, VPD proved to be a stronger sedative and neurotoxin than VPA in a mice model of seizures [68].

Reports show that valpromide gets partially converted to valproic acid in dogs and rats [71]. Consistent with this, Bailer et al. found that following the oral administration of VPD in human subjects, blood plasma detection of VPA was observed instead of VPD. This suggested that VPD acts as a prodrug to VPA in humans, and bio-transforms into VPA in humans, due to which its superiority over VPA in vivo does not have any clinical significance [72].

To build upon the reduced teratogenicity, improved potency and to resist metabolic hydrolysis of the amide to its parent carboxylic acid compound, a further three compounds were derived from VPD (valnoctamide, sec-butyl-propyl carboxamide and propylisopropyl acetamide) [56].

Valnoctamide (VCD, 2-Ethyl-3-methylpentanamide)

Valnoctamide is a 143.23 g/mol constitutional isomer of VPD with valnoctic acid (VCA) as its corresponding acid. Structurally, the parent chain has two branches instead of one making two chiral centres. The first branching is from an ethyl group at C-2 position and the second branching is from a methyl group at C-3 position. Similar to VPD, VCD has also shown higher potency than VPA in rodent models, which might be attributed to its better brain distribution [50]. In a pilocarpine-treated rat model of focal epilepsy, VCD proved as effective as its parent compound VPD and its analogue VPA in showing protective and antiepileptogenic effects [73], although it failed to show any activity in status epilepticus (SE) model [51]. Additionally, both VCD and VPD also demonstrated improved attenuation of extracellular glutamate levels as compared to VPA [73].

Apart from anticonvulsant action, the enantiomers of VCD have been demonstrated to show efficacious antiallodynic activity in rats with neuropathic pain induced by spinal nerve ligation (SNL) [74]. Of the two diastereoisomers, (2*S*,3*S*)-VCD demonstrated lesser embryotoxicity and higher potency than (2*R*,3*S*)-VCD, indicating that (2*S*,3*S*)-VCD as the potential molecule for neuropathic pain treatment [74].

VCD was found to have longer half-life and mean residence time than VPD [50]. Besides, unlike VPD, VCD does not bio-transform into its corresponding acid or VPA under physiological condition and thus has a greatly reduced teratogenic potential. Okada et al. reported that inhibition of Polycomb group (Pc-G) gene expression by VPA led to axial skeletal abnormalities in embryos of pregnant NMRI mice, whereas VPD and VCD did not show any significant effect on these genes, thus proving them to be much less teratogenic than VPA [69]. The same group also revealed that VPA affects expression changes in genes involved in the cell cycle and apoptosis pathways of neural tube cells, causing neural tube defects in embryos of NMRI mice. However, VCD and VPD affected only few of these genes with majority of genes unaffected [75]. VCD has undergone phase III clinical trials and is primarily prescribed for mania, bipolar, and schizoaffective disorders [66,74]. Nirvanil, which is a racemic mixture of both enantiomeric pairs of VCD, is a marketed anxiolytic drug in some European countries, namely France, Holland, Italy, and Switzerland [76,77].

sec-Butyl-propylacetamide (SPD, 3-Methyl-2-propylpentanamide)

Another amide of VPA developed from VCD is its one-carbon homologue namely *sec*-Butyl-propylacetamide (SPD) and has been under investigation as a replacement of VPA. SPD differs from VCD by having a propyl rather than ethyl side chain at C-2 position. SPD demonstrated broad-spectrum antiseizure profile comparable to VPA [51]. Being currently in phase II clinical trials, SPD in animal studies has been effective against both focal and secondarily generalized seizures [51]. A dose-dependent decline was observed in secondarily generalized seizures when SPD was tested on kindled rodent models with median effective dose (ED_{50}) values lesser than VCD and hence better efficacy [51]. It also caused a dip in seizure score from 4.6 to 1.3 in hippocampal kindled rats at a dose of 32 mg/kg and also with corneal kindled mice from 5 to 1 at 60 mg/kg [51]. In case of refractory 6Hz limbic seizures, an ED_{50} = 27 mg/kg for SPD was observed which is less than that of VCD (ED_{50} = 37 mg/kg) [51]. Moreover, unlike VCD, SPD also showed its activity in pilocarpine SE model with an ED_{50} value of 84 mg/kg with improved neuroprotective role as compared to VPA. SPD has also shown promising results in managing benzodiazepine-resistant seizures [51]. These findings implicate better efficacy of SPD over VCD.

Kaufmann et al. studied the antiallodynic activity of range of VPA related compounds in SNL model and found SPD to be as effective as Gabapentin in countering neuropathic pain with ED_{50} values of 49 mg/kg along with mild sedation observed at dose of 80 mg/kg [78].

The lack of teratogenicity was shown in studies where racemic SPD failed to induce neural tube defects (NTD) in SWV mice. Thus, along with better anticonvulsant property, it is nonembryotoxic and non-teratogenic [79].

Propylisopropyl Acetamide (PID, Diisopropyl Acetamide)

PID is a 143.23 g/mol constitutional isomer of VPD with a single stereocenter. As the name indicates, it has an isopropyl side chain at C-2 parent carbon instead of a n-propyl group as it is in VPA. It has anticonvulsant ability comparable to VPD and about 3–30 times more potency than VPA [80,81]. PID and VCD were tested for their anticonvulsant activity against generalized convulsive seizures in developing rats induced with PTZ. PID showed better control of generalized tonic–clonic seizures (GTCS) in 25 days old rats as compared to much younger ones, whereas VCD proved to be an overall better anticonvulsant, with no significant differences observed between the two stereoisomers [82]. Stereoselectivity of PID enantiomers was seen in animal models, where it was observed that the R enantiomer is a more potent anticonvulsant than S-PID with ED_{50} of R-PID being 19–32% lower than the latter [83]. It was also found to be more potent in 6Hz seizure model compared to S-PID [81].

The results of teratogenic effect of VPA and PID in pregnant NMRI mice exhibited that in contrast to VPA causing 37% and 73% (water and Cremophor EL suspension respectively) exencephaly in foetuses; PID and its enantiomers did not show any significant teratogenic effect [84].

The development of PID as second generation VPA was stopped due to financial limitations of the holding company Jazz Pharmaceuticals [85].

3.2. Acid Analogues of VPA

3.2.1. 2-ene-VPA (2-Propyl-2-pentenoic Acid)

2-ene-VPA is a 142.2 g/mol, unsaturated metabolite of VPA with a double bond between C-2 and C-3 has undergone phase I clinical trial in humans to establish its safety and tolerability [53]. Its potency was observed to be 60–90% of the parent compound in electro and chemiconvulsive threshold tests in vivo [86,87]. Loscher et al. did the comparative analysis of 2-ene-VPA against the parent compound in mice models with MES and PTZ seizure, rats with chronically reoccurring spontaneous petit mal seizures and gerbils with GTCS [88]. The results of the study clearly reflected that 2-ene-VPA has comparable ED_{50} values to that of VPA and no toxicity at effective dose in gerbils and rats. The study also reported lack of embryotoxicity even at doses as high as 600 mg/kg [88]. However, 2-ene-VPA showed greater sedation than VPA at a dose of 200–300 mg/kg in MES- and PTZ-induced seizures [89].

The pharmacokinetic shortcoming of this analogue was its partial biotransformation into VPA, due to which its development as an AED was ceased [90].

3.2.2. 4-ene-VPA (2-n-Propyl-4-pentenoic Acid)

4-ene-VPA is a 142.2 g/mol acidic analogue and metabolite formed by terminal desaturation of valproic acid, by the action of CYP2C9 [44]. It is shown to have anticonvulsant potency similar to VPA [86,87].

In a study comparing teratogenic potencies of R and S enantiomeric forms of 4-ene-VPA in NMRI mice, S-(-)-4-ene-VPA was revealed to show four times higher exencephaly compared to the R enantiomer and also had higher teratogenic potential than VPA itself [54]. 4-ene-VPA also demonstrated hepatotoxicity in rats, and proved to be a good prediction compound in urine to ascertain hepatotoxicity of VPA [27]. Thus, higher hepatoxicity and teratogenicity of the metabolite than VPA despite having similar anticonvulsant activity might reason the absence of its clinical trial record.

3.3. Fluorinated Derivatives

Alpha-fluorovalproic acid is a 162.2 g/mol fluorinated derivative of valproic acid. This derivatisation was introduced with an aim to reduce the hepatotoxicity of VPA [25]. As mentioned earlier, metabolic conversion of VPA into (E)-2,4-diene-VPA via CYP450 and mitochondrial beta-oxidation results in further toxic intermediates, which causes hepatotoxicity. The introduction of fluorine at the alpha carbon of VPA, forms α-fluoro-4-ene VPA upon metabolic conversion [91]. This intermediate is a non-toxic analogue of 4-ene

VPA and does not undergo beta-oxidation, thus preventing the formation of downstream hepato-toxic metabolite, 2,4-diene-VPA [92]. This non-toxic effect was evaluated in rats administered with α-fluoro-4-ene VPA, wherein no signs of hepatic microvesicular steatosis were observed. The derivative also showed reduced teratogenic effect in NMRI mice compared to VPA [61].

Although α-fluoro VPA did show its suppressive effect against PTZ-induced seizures in CD-1 mice, its maximal effect was much delayed likely due to its slower brain uptake [24]. Its ED_{50} value was observed to be 275.74 mg/kg vs. that of VPA with an ED_{50} of 119.7 mg/kg [24].

3.4. Cyclic Analogues of VPA

3.4.1. Cyclooctylideneacetic Acid (2-Cycloctylideneacetic Acid)

Cyclooctylideneacetic acid is a 168.23 g/mol cyclic acid analogue of VPA. In a study by Palaty J and Abbott F.S., mice with PTZ-induced seizure were subjected to 17 VPA-related cyclic and acyclic analogues. Of these 17 analogues, cyclooctylideneacetic acid was found to be the most potent anticonvulsant, even exceeding the potency of VPA with an ED_{50} dose (122.8 mg/kg) lower than VPA [55]. Moreover, the compound imparted a significantly reduced sedative effect on the treated mice at the observed ED_{50} value [55].

3.4.2. Tetramethylcyclopropyl Analogues

These analogues of VPA were designed to inhibit the conversion of VPA to its hepatotoxic metabolites [93]. The lack of potential transformation of this analogue is due to presence of two tertiary carbons at beta position to the carboxyl moiety [64].

2,2,3,3-Tetramethylcyclopropanecarboxylic Acid (TMCA)

TMCA is a 142 g/mol cyclopropyl analogue of VPA. It has weak anticonvulsant activity as compared to VPA, but has enhanced antiallodynic activity compared to parent drug, at doses that do not show any sedation effect [62]. This was shown in SNL mice model for tactile allodynia where the antiallodynic effect was observed at ED_{50} value of 181 mg/kg vs. ED_{50} value of 269 mg/kg of VPA 60 min post dosage [62].

In a fashion similar to the production of VPA amides with improved potency, TMCA amide derivatives were also studied as discussed in below sections.

2,2,3,3-Tetramethylcyclopropanecarboxamide (TMCD)

TMCD is a 141 g/mol cyclopropyl analogue of VPD. Okada et al. observed the analogue to have better anticonvulsant activity and lesser teratogenicity in mice embryos than VPA [64]. Consistent finding was also reported in rats, where reduced fetal skeletal abnormalities were observed as compared to VPA [63]. Hence, the analogue was suggested as a possible candidate for VPA replacement. Similar to TMCA, TMCD was also found to be potent antiallodynic drug with ED_{50} of 85 mg/kg [62].

N-Methoxy-TMCD (MTMCD)

It is a 154 g/mol derivative of TMCD, where the hydrogen atom of NH_2 in TMCD is substituted with a methoxy group. MTMCD is a broad-spectrum anticonvulsant [63,94]. In the same study by Okada et al., enhanced anticonvulsant activity for MTMCD in comparison to VPA was observed, accompanied with less teratogenicity shown in vivo [64]. This compound was found to have 18.5 times more potency than VPA in rat subcutaneous metrazol (scMet) test and 4.5 times in MES test. It also has a wider safety margin (8 times the protective index (PI = TD_{50}/ED_{50}) of VPA) [64]. Among other cyclopropyl analogues of VPA, MTMCD showed best antiallodynic activity with ED_{50} value of 41 mg/kg [62].

Alpha-Fluro-TMCD

As mentioned, fluorinated VPA analogues show reduced hepatotoxicity but were also accompanied by reduced anticonvulsant activity [24,25,92]; this limitation was overcome

by fluorination of TMCD [65]. Alpha-fluoro TMCD showed 120 times more potent anticonvulsant activity in scMET rat model than VPA [65]. It also showed higher potency in other in vivo test (kindled rat model, 6 Hz test and pilocarpine induced seizure in rat) and was found to be non-teratogenic with better safety margin than VPA, due to which it was proposed to be a potent candidate of CNS therapeutics [65].

Other compounds in this category are cyclohexanecarboxylic acid and its methylated form. Cyclohexanecarboxylic acid, a 128.17 g/mol analogue, showed promise with reduced neurotoxicity, though it still maintains the low potency of VPA [90]. In contrast, the 142.2 g/mol methylated analogue, i.e., 1-methylcyclohexane carboxylic acid depicted higher potency, but had a fatally increased neurotoxic potential [90].

3.5. Urea Derivatives of VPA

In the studies targeted at understanding the structure–activity relationships between VPA and its adverse effects, a number of urea derivatives have also been synthesized. The three compounds showing promise as future anticonvulsants are derived from VPA amide derivatives, and have not only improved anticonvulsant activity but also broader spectrum [56].

3.5.1. Valnoctylurea (VCU, 2-Ethyl-3-methylpentanoyl Urea)

VCU is a 186 g/mol urea derivative of VCA. VCU along with diisopropyl acetic acid (DIU) has shown tremendous anticonvulsant potential in scMET and MES seizure model, even better than amide derivatives of VPA [56]. The observed ED_{50} values for VCU are 14 mg/kg and 48 mg/kg in scMET seizures and 6Hz seizures, respectively, which was much better than those observed for VPA (646 mg/kg and 310 mg/kg respectively) [56]. In addition, VCU also has a lower teratogenic effect, leading to its acceptance as a safer derivative of VPA [56].

3.5.2. Propyl Isopropylacetyl Urea (PIU)

PIU is a 186 g/mol urea derivative of PID with the characteristic C-2 chiral centre. The enantioselective anticonvulsant effect of (R) and (S)-PIU was consistently observed in both MES and scMET-induced seizures [56]. In MES model, an ED_{50} value of 16 mg/kg was determined for the racemic mixture of PIU enantiomers, whereas individual R-PIU and S-PIU enantiomers had ED_{50} 36 mg/kg and 18 mg/kg respectively [56]. Opposite results were observed in scMET model, where (R)-PIU had better efficacy with ED_{50} of 22 mg/kg that was significantly lower than S-PIU (ED_{50} = 37 mg/kg) and racemic-PIU (ED_{50} = 45 mg/kg) [56]. Results similar to scMET model was also observed for 6Hz model.

Teratogenic effect of (R)-PIU was observed at doses 5 to 6 times higher than the its ED_{50} [56]. This effect was greater than the S enantiomer. The significant difference in the toxicity of the two enantiomers can be ascertained as (R)-PIU has 54% embryotoxicity as compared to 10% of (S)-PIU [56]. The R isomer also leads in causing neural tube defects in the exposed embryos [56]. Yet the safety margin of both the enantiomers was better than VPA at their effective doses.

3.5.3. Diisopropyl Acetyl Urea (DIU)

DIU is a 186 g/mol urea derivative of diisopropyl acetyl amide (DIA). DIU has shown itself to be a promising anticonvulsant against 6Hz seizures with ED50 of 49 mg/kg that is sufficiently lower than VPA (310 mg/kg) [56]. It also showed better results than VPA in scMET and MES model with ED_{50} value of 16 mg/kg and 33 mg/kg, respectively, compared to ED_{50} value of 646 mg/kg and 485 mg/kg of VPA. Like other urea derivatives of VPA amide, DIU has better safety margin than VPA [56].

3.6. Conjugation Products of VPA

Valrocemide (TV1901, VGD)

VGD is a 200.28 g/mol amide derivative conjugation product of VPA and glycinamide. The conjugate was targeted to maximize the brain penetration [95]. Drug-distribution study performed by Blotnik et al. documented 4 times higher brain-to-plasma ratio of VGD than VPA in rats, thus implying better brain distribution of the conjugate [96]. VGD turned out to be a VPA-like broad-spectrum anticonvulsant [57]. VGD has proven itself to show doubly potent anticonvulsant activity than VPA and can effectively counter MES, 6-Hz psychomotor seizures (even at 44 mA stimulation), PTZ-, picrotoxin-, and bicuculline-induced seizures in rat model. VGD lagged behind VPA only in managing seizures of hippocampal kindled rats. The compound has reached Phase II clinical trials as an AED for patients with refractory epilepsy [97].

VGD also lacks a carboxylic moiety, thus its teratogenic potential is essentially missing and, therefore, no teratogenic effect of the compound was observed in rats, rabbits or SWV mice [57].

3.7. Prodrugs (Sugar Esters of VPA)

Unlike the other compounds discussed above, the sugar esters do not aim to change the VPA-related adverse drug reactions (ADRs). Sugar esters as prodrugs of VPA were developed as slow-release formulations of VPA to counter its limited plasma half-life, though they were also found to exert direct action on the brain. Amongst these esters, dimethylenexylitol valproate (VDMX) turned up as the most potent compound with its effective dose as low as 100 times that of VPA in suppressing spontaneous epileptiform activities (SEA) in rat hippocampal slices [60]. Also, VDMX showed best protective index in PTZ test [60]. VDMX shows good activity in entorhinal cortex late recurrent discharges (LRD) model for pharmacoresistant epilepsy, even though VPA fails to exert any significant effect on suppressing epileptiform activity in LRDs [59]. It was speculated that since LRD primarily depend on glutamatergic mechanisms, VDMX might strongly affect glutamatergic synaptic transmissions to exert its actions.

4. Conclusions

Introduced in clinical practice about 60 years ago, valproic acid still represents one of the most efficient AED targeting broad spectrum of epilepsy types. Emerging evidence also demonstrates its use in other neurological disorders such as migraine, bipolar disorder and neuropathic pain. However, its use in pregnant women with epilepsy (WWE) has shown to be associated with teratogenic effects on the developing foetus leading to congenital abnormalities. In addition, hepatotoxicity is another adverse effect commonly experienced with VPA therapy.

Numerous attempts have been made by various research groups to modify VPA structure. These were focused on altering one of the factors deemed responsible for its teratogenic behaviour or on avoiding the formation of its hepatotoxic metabolites (4-ene-VPA and 2,4-diene-VPA) while simultaneously increasing or maintaining its anticonvulsant activity and potency. This review has meticulously accounted for all the structural derivatives and analogues of VPA, discussing their pharmacological and toxicological profiles tested in vivo. These included amide derivatives, acid analogues, fluorinated derivatives, cyclic analogues, urea derivatives, conjugation products and sugar esters. The modifications thus made, have proven themselves to be effective as anxiolytics, antiallodynics and better anticonvulsants than VPA. Furthermore, the majority of them have shown better potency and tolerability in vivo.

Among the developed amide derivatives, VCD and SPD were found to have potentially no teratogenicity, are broad spectrum AEDs just like VPA, and are under active investigations. VCD has shown itself to be non-teratogenic at therapeutic doses in animal models [98,99] and has an anticonvulsant profile comparable to SPD. Advantages of SPD over VCD are lower ED_{50} values in pilocarpine induced status epilepticus [66,100,101].

Even though the fluorinated derivatives of VPA have reduced hepatotoxicity and teratogenicity, they are not as effective anticonvulsants as VPA, thus stopping their development. A fluorinated hydroxamic acid derivative of; VPA 2-Fluoro-VPA-HA [61] and alpha-Fluoro-TMCD [65] proved to be good candidates as second generation VPA. Apart from this, effectiveness of VGD in controlling mouse 44mA 6Hz-induced seizures and its better brain distribution has led to its successful completion of a 13-week, phase II clinical trial in Europe with resistant patients with epilepsy [57,58,77].

This article provides necessary knowledgebase of existing modifications which may facilitate more rational attempts in making structural changes. These strategies will systematically enhance the therapeutic use of VPA-related compounds for better disease management.

Author Contributions: R.K.: conceptualization and supervision. M.K.M., S.K. (Samiksha Kukal), P.R.P. and S.B.: original draft preparation. S.K. (Samiksha Kukal) and P.R.P.: editing. A.S., S.K. (Shrikant Kukreti), L.S., K.M. and Y.H.: manuscript revision. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by Council of Scientific and Industrial Research (CSIR) under Grant OLP1154 and OLP1142.

Acknowledgments: We are thankful to the Director, CSIR-Institute of Genomics and Integrative Biology (IGIB), Anurag Agrawal, for his scientific vision and support. We are also thankful to the research scholar, Neha Kanojia, CSIR-IGIB, for her help in literature screening for the review. M.K.M. acknowledge Indian Council of Medical Research (ICMR), Govt. of India, S.K. (Samiksha Kukal) acknowledge Department of Biotechnology (DBT) and CSIR, Govt. of India, P.R.P. and S.B. acknowledge CSIR, Government of India, for their financial assistance. We thank the anonymous reviewers for their helpful suggestions in improving the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

VPA: valproic acid: CNS, central nervous system; AED, antiepileptic drug; GABA, gamma-aminobutyric acid; OAT, organic anion transporters; CYP450, cytochrome P450; GABA-T, GABA transaminase; SSADH, succinate semialdehyde dehydrogenase; GluR1/2, glutamate receptor subunits 1 and 2; SCN3A, sodium voltage-gated channel alpha subunit 3; HDAC, histone deacetylase; GABAR: GABA receptors; BDNF: brain derived neurotrophic factor; CCND2: cyclin D2; PGE2: prostaglandin E2; GSK3B: glycogen synthase kinase 3 beta; PKC: protein kinase C; ERK/BCL-2: extracellular regulated kinase—B-cell lymphoma 2 pathway; Bad: BCL2 associated agonist of cell death; CDK5: cyclin-dependent kinase 5; NEP: neprilysin; TGF-B: transforming growth factor beta; VEGF: vascular endothelial growth factor, IL-1B: interleukin 1 beta; CCR1: C-C chemokine receptor type 1; p-JNK: phosphorylated c-Jun N-terminal kinase; H3K9ac: histone 3 lysine 9 acetylation; Ngb: neuroglobin; Nrf2/HO-1: nuclear factor erythroid 2–related factor 2—heme oxygenase-1 pathway; H3ac: histone 3 acetylation; H4ac: histone 4 acetylation; p21: cyclin-dependent kinase inhibitor; MYC: master regulator of cell cycle entry and proliferative metabolism; CCND1: cyclin D1; NOTCH1: Notch homolog 1, translocation-associated; p63: tumor protein p63; IL-6: interleukin 6; ANF: atrial natriuretic peptide; MES, maximal electroshock stimulation; PTZ, pentylenetetrazole; VPD, valpromide; SOX9, SRY-Box Transcription Factor 9; RUNX2, Runt-related transcription factor 2; VCD, valnoctamide, VCA, valnoctic acid; SE, status epilepticus; SNL, spinal nerve ligation; Pc-G, Polycomb group; NMRI, Naval Medical Research Institute; SPD, sec-Butyl-propylacetamide; ED_{50}, effective dose; NTD, neural tube defect; PID, propylisopropyl acetamide; GTCS, generalized tonic–clonic seizures; TMCA, 2,2,3,3-tetramethylcyclopropanecarboxylic acid; TMCD, 2,2,3,3-tetramethylcyclopropanecarboxamide; MTMCD, N-methoxy-TMCD; scMET, subcutaneous Metrazol; VCU, valnoctylurea; DIU, diisopropyl acetyl urea; DIA, diisopropyl acetic acid. PIU, propyl isopropylacetyl urea; VGD, valrocemide; ADR, adverse drug reaction; SEA, spontaneous epileptiform activities; VDMX, dimethylenexylitol valproate; LRD, late recurrent discharges; WWE, women with epilepsy.

References

1. Perucca, E. Pharmacological and Therapeutic Properties of Valproate: A Summary after 35 Years of Clinical Experience. *CNS Drugs* **2002**, *16*, 695–714. Available online: https://pubmed.ncbi.nlm.nih.gov/12269862/ (accessed on 25 November 2021). [CrossRef]
2. Ghodke-Puranik, Y.; Thorn, C.F.; Lamba, J.K.; Leeder, J.S.; Song, W.; Birnbaum, A.K.; Altman, R.B.; Klein, T.E. Valproic Acid Pathway: Pharmacokinetics and Pharmacodynamics. *Pharm. Genom.* **2013**, *23*, 236–241. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3696515/ (accessed on 29 September 2020). [CrossRef]
3. López-Muñoz, F.; Baumeister, A.A.; Hawkins, M.F.; Álamo, C. The Role of Serendipity in the Discovery of the Clinical Effects of Psychotropic Drugs: Beyond of the Myth. *Actas Esp. Psiquiatr.* **2012**, *40*, 34–42.
4. Meunier, H.; Carraz, G.; Neunier, Y.; Eymard, P.; Aimard, M. Pharmacodynamic Properties of N-dipropylacetic acid. *Therapie* **1963**, *18*, 435–438. Available online: http://europepmc.org/article/med/13935231 (accessed on 29 September 2020).
5. Diederich, M.; Chateauvieux, S.; Morceau, F.; Dicato, M. Molecular and Therapeutic Potential and Toxicity of Valproic Acid. *J. Biomed. Biotechnol.* **2010**. Available online: https://pubmed.ncbi.nlm.nih.gov/20798865/ (accessed on 29 September 2020).
6. Emrich, H.M.; Von Zerssen, D.; Kissling, W.; Moeller, H.J. Therapeutic Effect of Valproate in Mania. *Am. J. Psychiatry* **1981**, *138*, 256. Available online: https://pubmed.ncbi.nlm.nih.gov/6779643/ (accessed on 29 September 2020).
7. Calabresi, P.; Galletti, F.; Rossi, C.; Sarchielli, P.; Cupini, L.M. Antiepileptic drugs in Migraine: From Clinical Aspects to Cellular Mechanisms. *Trends Pharmacol. Sci.* **2007**, *28*, 188–195. Available online: https://linkinghub.elsevier.com/retrieve/pii/S0165614707000491 (accessed on 29 September 2020). [CrossRef]
8. Tariot, P.N.; Loy, R.; Ryan, J.M.; Porsteinsson, A.; Ismail, S. Mood Stabilizers in Alzheimer's Disease: Symptomatic and Neuroprotective Rationales. *Adv. Drug Deliv. Rev.* **2002**, *54*, 1567–1577. Available online: https://linkinghub.elsevier.com/retrieve/pii/S0169409$\times$02001539 (accessed on 29 September 2020). [CrossRef]
9. Loy, R.; Tariot, P.N. Neuroprotective Properties of Valproate: Potential Benefit for AD and Tauopathies. *J. Mol. Neurosci.* **2002**, *19*, 303–307. Available online: https://pubmed.ncbi.nlm.nih.gov/12540056/ (accessed on 25 November 2021). [CrossRef]
10. Nielsen, N.M.; Svanström, H.; Stenager, E.; Magyari, M.; Koch-Henriksen, N.; Pasternak, B.; Hviid, A. The Use of Valproic Acid and Multiple Sclerosis. *Pharmacoepidemiol. Drug Saf.* **2015**, *24*, 262–268. Available online: https://pubmed.ncbi.nlm.nih.gov/25111895/ (accessed on 25 November 2021). [CrossRef] [PubMed]
11. Blaheta, R.; Nau, H.; Michaelis, M.; Cinatl, J., Jr. Valproate and Valproate-Analogues: Potent Tools to Fight Against Cancer. *Curr. Med. Chem.* **2012**, *9*, 1417–1433. Available online: https://pubmed.ncbi.nlm.nih.gov/12173980/ (accessed on 29 September 2020). [CrossRef]
12. Phiel, C.J.; Zhang, F.; Huang, E.Y.; Guenther, M.G.; Lazar, M.A.; Klein, P.S. Histone Deacetylase is a Direct Target of Valproic Acid, a Potent Anticonvulsant, Mood Stabilizer, and Teratogen. *J. Biol. Chem.* **2001**, *276*, 36734–36741. Available online: https://pubmed.ncbi.nlm.nih.gov/11473107/ (accessed on 10 September 2020). [CrossRef]
13. Göttlicher, M.; Minucci, S.; Zhu, P.; Krämer, O.H.; Schimpf, A.; Giavara, S.; Sleeman, J.P.; Coco, F.L.; Nervi, C.; Pelicci, P.G.; et al. Valproic Acid Defines a Novel Class of HDAC Inhibitors inducing Differentiation of Transformed Cells. *EMBO J.* **2001**, *20*, 6969–6978. Available online: https://pubmed.ncbi.nlm.nih.gov/11742974/ (accessed on 22 November 2021). [CrossRef]
14. Bradbury, C.A.; Khanim, F.L.; Hayden, R.; Bunce, C.M.; White, D.A.; Drayson, M.T.; Craddock, C.; Turner, B.M. Histone Deacetylases in Acute Myeloid Leukaemia Show a Distinctive Pattern of Expression that Changes Selectively in Response to Deacetylase Inhibitors. *Leukemia* **2005**, *19*, 1751–1759. Available online: https://pubmed.ncbi.nlm.nih.gov/16121216/ (accessed on 22 November 2021). [CrossRef] [PubMed]
15. Jäger-Roman, E.; Deichl, A.; Jakob, S.; Hartmann, A.M.; Koch, S.; Rating, D.; Steldinger, R.; Nau, H.; Helge, H. Fetal Growth, Major Malformations, and Minor Anomalies in Infants Born to Women Receiving Valproic Acid. *J. Pediatr.* **1986**, *108*, 997–1004. Available online: https://pubmed.ncbi.nlm.nih.gov/3086531/ (accessed on 30 September 2020). [CrossRef]
16. Terbach, N.; Shah, R.; Kelemen, R.; Klein, P.S.; Gordienko, D.; Brown, N.A.; Wilkinson, C.J.; Williams, R.S. Identifying an Uptake Mechanism for the Antiepileptic and Bipolar Disorder Treatment Valproic acid Using the Simple Biomedical Model Dictyostelium. *J. Cell Sci.* **2011**, *124 Pt 13*, 2267–2276. Available online: https://pubmed.ncbi.nlm.nih.gov/21652627/ (accessed on 22 November 2021). [CrossRef] [PubMed]
17. Keane, P.E.; Simiand, J.; Mendes, E.; Santucci, V.; Morre, M. The Effects of Analogues of Valproic Acid on Seizures Induced by Pentylenetetrazol and GABA Content in Brain of Mice. *Neuropharmacology* **1983**, *22*, 875–879. Available online: https://pubmed.ncbi.nlm.nih.gov/6413882/ (accessed on 22 November 2021). [CrossRef]
18. Chapman, A.G.; Meldrum, B.S.; Mendes, E. Acute Anticonvulsant Activity of Structural Analogues of Valproic Acid and Changes in Brain GABA and Aspartate Content. *Life Sci.* **1983**, *32*, 2023–2031. Available online: https://pubmed.ncbi.nlm.nih.gov/6403794/ (accessed on 22 November 2021). [CrossRef]
19. Chapman, A.G.; Croucher, M.J.; Meldrum, B.S. Anticonvulsant Activity of Intracerebroventricularly Administered Valproate and Valproate Analogues. A Dose-Dependent Correlation with Changes in Brain Aspartate and GABA Levels in DBA/2 Mice. *Biochem. Pharmacol.* **1984**, *33*, 1459–1463. Available online: https://pubmed.ncbi.nlm.nih.gov/6428419/ (accessed on 22 November 2021). [CrossRef]
20. Comelli, N.C.; Duchowicz, P.R.; Lobayan, R.M.; Jubert, A.H.; Castro, E.A. QSPR Study of Valproic Acid and Its Functionalized Derivatives. *Mol. Inform.* **2012**, *31*, 181–188. Available online: https://pubmed.ncbi.nlm.nih.gov/27476963/ (accessed on 25 November 2021). [CrossRef] [PubMed]

21. Nau, H.; Hauck, R.-S.; Ehlers, K. Valproic Acid-Induced Neural Tube Defects in Mouse and Human: Aspects of Chirality, Alternative Drug Development, Pharmacokinetics and Possible Mechanisms. *Pharmacol. Toxicol.* **1991**, *69*, 310–321. Available online: https://pubmed.ncbi.nlm.nih.gov/1803343/ (accessed on 22 November 2021). [CrossRef]
22. Nau, H.; Löscher, W. Pharmacologic Evaluation of Various Metabolites and Analogs of Valproic Acid: Teratogenic Potencies in mice. *Fundam. Appl. Toxicol.* **1986**, *6*, 669–676. Available online: https://pubmed.ncbi.nlm.nih.gov/3086174/ (accessed on 22 November 2021). [CrossRef]
23. Lloyd, K.A.; Sills, G. A Scientific Review: Mechanisms of Valproate-Mediated Teratogenesis. *Biosci. Horiz. Int. J. Stud. Res.* **2013**, *6*, 2013. Available online: https://academic.oup.com/biohorizons/article/doi/10.1093/biohorizons/hzt003/302011 (accessed on 25 November 2021). [CrossRef]
24. Tang, W.; Palaty, J.; Abbott, F.S. Time Course of Alpha-Fluorinated Valproic Acid in Mouse Brain and Serum and Its Effect on Synaptosomal Gamma-Aminobutyric Acid Levels in Comparison to Valproic Acid. *J. Pharmacol. Exp. Ther.* **1997**. Available online: https://pubmed.ncbi.nlm.nih.gov/9316822/ (accessed on 22 November 2021).
25. Tang, W.; Borel, A.G.; Abbott, F.S.; Fujimiya, T. Fluorinated Analogues as Mechanistic Probes in Valproic Acid Hepatotoxicity: Hepatic Microvesicular Steatosis and Glutathione Status. *Chem. Res. Toxicol.* **1995**, *8*, 671–682. Available online: https://pubmed.ncbi.nlm.nih.gov/7548749/ (accessed on 22 November 2021). [CrossRef]
26. Neuman, M.G.; Shear, N.H.; Jacobson-Brown, P.M.; Katz, G.G.; Neilson, H.K.; Malkiewicz, I.M.; Cameron, R.G.; Abbott, F. CYP2E1-Mediated Modulation of Valproic Acid-Induced Hepatocytotoxicity. *Clin. Biochem.* **2001**, *34*, 211–218. Available online: https://pubmed.ncbi.nlm.nih.gov/11408019/ (accessed on 22 November 2021). [CrossRef]
27. Lee, M.S.; Lee, Y.J.; Kim, B.J.; Shin, K.J.; Chung, B.C.; Baek, D.J.; Jung, B.H. The Relationship between Glucuronide Conjugate Levels and Hepatotoxicity after Oral Administration of Valproic Acid. *Arch Pharm. Res.* **2009**, *32*, 1029–1035. Available online: https://pubmed.ncbi.nlm.nih.gov/19641884/ (accessed on 10 September 2020). [CrossRef]
28. Zaccara, G.; Messori, A.; Moroni, F. Clinical Pharmacokinetics of Valproic Acid—1988. *Clin. Pharmacokinet.* **1988**, *15*, 367–389. Available online: https://pubmed.ncbi.nlm.nih.gov/3149565/ (accessed on 22 November 2021). [CrossRef]
29. Romoli, M.; Mazzocchetti, P.; D'Alonzo, R.; Siliquini, S.; Rinaldi, V.E.; Verrotti, A.; Calabresi, P.; Costa, C. Valproic Acid and Epilepsy: From Molecular Mechanisms to Clinical Evidences. *Curr. Neuropharmacol.* **2019**, *17*, 926–946. Available online: https://pubmed.ncbi.nlm.nih.gov/30592252/ (accessed on 22 November 2021). [CrossRef]
30. Levy, R.H.; Cenraud, B.; Loiseau, P.; Akbaraly, R.; Brachet-Liermain, A.; Guyot, M.; Gomeni, R.; Morselli, P.L. Meal-Dependent Absorption of Enteric-Coated Sodium Valproate. *Epilepsia* **1980**, *21*, 273–280. Available online: https://pubmed.ncbi.nlm.nih.gov/6769666/ (accessed on 22 November 2021). [CrossRef]
31. Ito, M.; Ikeda, Y.; Arnez, J.G.; Finocchiaro, G.; Tanaka, K. The Enzymatic Basis for the Metabolism and Inhibitory Effects of Valproic Acid: Dehydrogenation of valproyl-CoA by 2-methyl-branched-chain acyl-CoA Dehydrogenase. *Biochim. Biophys. Acta* **1990**, *1034*, 213–218. Available online: https://pubmed.ncbi.nlm.nih.gov/2112956/ (accessed on 22 November 2021). [CrossRef]
32. Argikar, U.A.; Remmel, R.P. Effect of Aging on Glucuronidation of Valproic acid in Human Liver Microsomes and the Role of UDP-glucuronosyltransferase UGT1A4, UGT1A8, and UGT1A10. *Drug Metab. Dispos.* **2009**, *37*, 229–236. Available online: https://pubmed.ncbi.nlm.nih.gov/18838507/ (accessed on 22 November 2021). [CrossRef]
33. Tan, L.; Yu, J.T.; Sun, Y.P.; Ou, J.R.; Song, J.H.; Yu, Y. The influence of cytochrome oxidase CYP2A6, CYP2B6, and CYP2C9 Polymorphisms on the Plasma Concentrations of Valproic Acid in Epileptic Patients. *Clin. Neurol. Neurosurg.* **2010**, *112*, 320–323. Available online: https://pubmed.ncbi.nlm.nih.gov/20089352/ (accessed on 22 November 2021). [CrossRef]
34. Rettie, A.E.; Rettenmeier, A.W.; Howald, W.N.; Baillie, T.A. Cytochrome P-450—Catalyzed Formation of Delta 4-VPA, a Toxic Metabolite of Valproic Acid. *Science* **1987**, *235*, 890–893. Available online: https://pubmed.ncbi.nlm.nih.gov/3101178/ (accessed on 22 November 2021). [CrossRef] [PubMed]
35. Bialer, M.; Hussein, Z.; Raz, I.; Abramsky, O.; Herishanu, Y.; Pachys, F. Pharmacokinetics of Valproic Acid in Volunteers after a Single Dose Study. *Biopharm. Drug Dispos.* **1985**, *6*, 33–42. Available online: https://pubmed.ncbi.nlm.nih.gov/3921078/ (accessed on 22 November 2021). [CrossRef]
36. Metabolite Pattern of Valproic Acid. Part I: Gaschromatographic Determination of the Valproic Acid Metabolite Artifacts, Heptanone-3, 4- and 5-Hydroxyvalproic Acid Lactone-PubMed. Available online: https://pubmed.ncbi.nlm.nih.gov/373770/ (accessed on 22 November 2021).
37. Identification and Characterization of the Glutathione and N-Acetylcysteine Conjugates of (E)-2-propyl-2,4-pentadienoic Acid, a Toxic Metabolite of Valproic Acid, in Rats and Humans-PubMed. Available online: https://pubmed.ncbi.nlm.nih.gov/1676665/ (accessed on 22 November 2021).
38. Kassahun, K.; Hu, P.; Grillo, M.P.; Davis, M.R.; Jin, L.; Baillie, T.A. Metabolic activation of unsaturated derivatives of valproic acid. Identification of novel glutathione adducts formed through coenzyme A-dependent and -independent processes. *Chem. Biol. Interact.* **1994**, *90*, 253–275. Available online: https://pubmed.ncbi.nlm.nih.gov/8168173/ (accessed on 23 November 2021). [CrossRef]
39. Rettenmeier, A.W.; Prickett, K.S.; Gordon, W.P.; Bjorge, S.M.; Chang, S.L.; Levy, R.H.; Baillie, T.A. Studies on the Biotransformation in the Perfused Rat Liver of 2-n-Propyl-4-pentenoic Acid, a Metabolite of the Antiepileptic Drug Valproic Acid. Evidence for the Formation of Chemically Reactive Intermediates-PubMed. Available online: https://pubmed.ncbi.nlm.nih.gov/2858383/ (accessed on 22 November 2021).

40. Baillie, T.A. Metabolic Activation of Valproic Acid and Drug-Mediated Hepatotoxicity. Role of the Terminal Olefin, 2-n-propyl-4-pentenoic Acid. *Chem. Res. Toxicol.* **1988**, *1*, 195–199. Available online: https://pubmed.ncbi.nlm.nih.gov/2979731/ (accessed on 22 November 2021). [CrossRef] [PubMed]
41. Peterson, G.M.; Naunton, M. Valproate: A Simple Chemical with so Much to Offer. *J. Clin. Pharm. Ther.* **2005**, *30*, 417–421. Available online: https://pubmed.ncbi.nlm.nih.gov/16164485/ (accessed on 22 November 2021). [CrossRef]
42. Riva, R.; Albani, F.; Contin, M.; Baruzzi, A.; Altomare, M.; Merlini, G.P.; Perucca, E. Mechanism of Altered Drug Binding to Serum Proteins in Pregnant Women: Studies with Valproic Acid. *Ther. Drug Monit.* **1984**, *6*, 25–30. Available online: https://pubmed.ncbi.nlm.nih.gov/6424276/ (accessed on 22 November 2021). [CrossRef]
43. Svensson, C.K.; Woodruff, M.N.; Baxter, J.G.; Lalka, D. Free Drug Concentration Monitoring in Clinical Practice. Rationale and Current Status. *Clin. Pharmacokinet.* **1986**, *11*, 450–469. Available online: https://pubmed.ncbi.nlm.nih.gov/3542337/ (accessed on 22 November 2021). [CrossRef] [PubMed]
44. Kiang, T.K.L.; Ho, P.C.; Anari, M.R.; Tong, V.; Abbott, F.S.; Chang, T.K.H. Contribution of CYP2C9, CYP2A6, and CYP2B6 to Valproic Acid Metabolism in Hepatic Microsomes from Individuals with the CYP2C9*1/*1 Genotype. *Toxicol. Sci.* **2006**, *94*, 261–271. Available online: https://pubmed.ncbi.nlm.nih.gov/16945988/ (accessed on 10 September 2020). [CrossRef] [PubMed]
45. VanDongen, A.M.J.; VanErp, M.G.; Voskuyl, R.A. Valproate Reduces Excitability by Blockage of Sodium and Potassium Conductance. *Epilepsia* **1986**, *27*, 177–182. Available online: https://pubmed.ncbi.nlm.nih.gov/3084227/ (accessed on 22 November 2021). [CrossRef] [PubMed]
46. Johannessen, C.U.; Johannessen, S.I. Valproate: Past, Present, and Future. *CNS Drug Rev.* **2003**, *9*, 199–216. Available online: https://pubmed.ncbi.nlm.nih.gov/12847559/ (accessed on 22 November 2021). [CrossRef]
47. Van den Berg, R.J.; Kok, P.; Voskuyl, R.A. Valproate and Sodium Currents in Cultured Hippocampal Neurons. *Exp. Brain Res.* **1993**, *93*, 279–287. Available online: https://pubmed.ncbi.nlm.nih.gov/8387930/ (accessed on 22 November 2021). [CrossRef]
48. Tan, N.N.; Tang, H.L.; Lin, G.W.; Chen, Y.H.; Lu, P.; Li, H.J.; Gao, M.M.; Zhao, Q.H.; Yi, Y.H.; Liao, W.P.; et al. Epigenetic Downregulation of Scn3a Expression by Valproate: A Possible Role in Its Anticonvulsant Activity. *Mol. Neurobiol.* **2017**, *154*, 2831–2842. Available online: https://pubmed.ncbi.nlm.nih.gov/27013471/ (accessed on 25 November 2021). [CrossRef]
49. Bialer, M.; Johannessen, S.I.; Kupferberg, H.J.; Levy, R.H.; Perucca, E.; Tomson, T. Progress Report on New Antiepileptic Drugs: A Summary of the Seventh Eilat Conference (EILAT VII). *Epilepsy Res.* **2004**, *61*, 1–48. Available online: https://pubmed.ncbi.nlm.nih.gov/15570674/ (accessed on 22 November 2021). [CrossRef] [PubMed]
50. Blotnik, S.; Bergman, F.; Bialer, M. Disposition of Valpromide, Valproic Acid, and Valnoctamide in the Brain, Liver, Plasma, and Urine of Rats-PubMed. Available online: https://pubmed.ncbi.nlm.nih.gov/8723737/ (accessed on 22 November 2021).
51. White, H.S.; Alex, A.B.; Pollock, A.; Hen, N.; Shekh-Ahmad, T.; Wilcox, K.S.; McDonough, J.H.; Stables, J.P.; Kaufmann, D.; Yagen, B.; et al. A New Derivative of Valproic Acid Amide Possesses a Broad-Spectrum Antiseizure Profile and Unique Activity against Status Epilepticus and Organophosphate Neuronal Damage. *Epilepsia* **2012**, *53*, 134–146. Available online: https://pubmed.ncbi.nlm.nih.gov/22150444/ (accessed on 10 September 2020). [CrossRef]
52. Bialer, M.; Johannessen, S.I.; Kupferberg, H.J.; Levy, R.H.; Perucca, E.; Tomson, T. Progress Report on New Antiepileptic Drugs: A Summary of the Eigth Eilat Conference (EILAT VIII). *Epilepsy Res.* **2007**, *73*, 1–52. Available online: http://www.ncbi.nlm.nih.gov/pubmed/17158031 (accessed on 22 November 2021). [CrossRef] [PubMed]
53. Düsing, R.H. Single-Dose Tolerance and Pharmacokinetics of 2-n-propyl-2(E)-pentenoate (delta 2(E)-valproate) in Healthy Male Volunteers. *Pharm. Weekbl. Sci.* **1992**, *14*, 152–158. Available online: https://pubmed.ncbi.nlm.nih.gov/1502017/ (accessed on 22 November 2021). [CrossRef]
54. Hauck, R.S.; Nau, H. Asymmetric Synthesis and Enantioselective Teratogenicity of 2-n-propyl-4-pentenoic Acid (4-en-VPA), an Active Metabolite of the Anticonvulsant Drug, Valproic Acid. *Toxicol. Lett.* **1989**, *49*, 41–48. Available online: https://pubmed.ncbi.nlm.nih.gov/2510370/ (accessed on 10 September 2020). [CrossRef]
55. Palaty, J.; Abbott, F.S. Structure-Activity Relationships of Unsaturated Analogues of Valproic Acid. *J. Med. Chem.* **1995**, *38*, 3398–3406. Available online: https://pubmed.ncbi.nlm.nih.gov/7650693/ (accessed on 18 November 2021). [CrossRef]
56. Shimshoni, J.A.; Bialer, M.; Wlodarczyk, B.; Finnell, R.H.; Yagen, B. Potent Anticonvulsant Urea Derivatives of Constitutional Isomers of Valproic Acid. *J. Med. Chem.* **2007**, *50*, 6419–6427. Available online: https://pubmed.ncbi.nlm.nih.gov/17994680/ (accessed on 10 September 2020). [CrossRef] [PubMed]
57. Isoherranen, N.; Woodhead, J.H.; White, H.S.; Bialer, M. Anticonvulsant Profile of Valrocemide (TV1901): A New Antiepileptic Drug. *Epilepsia* **2001**, *42*, 831–836. [CrossRef]
58. O'Shea, S.; Johnson, K.; Clark, R.; Sliwkowski, M.X.; Erickson, S.L. Effects of In Vivo Heregulin Beta1 Treatment in Wild-Type and ErbB Gene-Targeted Mice Depend on Receptor Levels and Pregnancy. *Am. J. Pathol.* **2001**, *158*, 1871–1880. Available online: http://www.ncbi.nlm.nih.gov/pubmed/11337386 (accessed on 10 September 2020). [CrossRef]
59. Armand, V.; Louvel, J.; Pumain, R.; Heinemann, U. Effects of New Valproate Derivatives on Epileptiform Discharges Induced by Pentylenetetrazole or Low Mg2+ in Rat Entorhinal Cortex-Hippocampus Slices. *Epilepsy Res.* **1998**, *32*, 345–355. Available online: https://pubmed.ncbi.nlm.nih.gov/9839774/ (accessed on 22 November 2021). [CrossRef]
60. Armand, V.; Louvel, J.; Pumain, R.; Ronco, G.; Villa, P. Effects of Various Valproic Acid Derivatives on Low-Calcium Spontaneous Epileptiform Activity in Hippocampal Slices. *Epilepsy Res.* **1995**, *22*, 185–192. Available online: https://pubmed.ncbi.nlm.nih.gov/8991785/ (accessed on 22 November 2021). [CrossRef]

61. Gravemann, U.; Volland, J.; Nau, H. Hydroxamic acid and Fluorinated Derivatives of Valproic Acid: Anticonvulsant Activity, Neurotoxicity and Teratogenicity. *Neurotoxicol. Teratol.* **2008**, *30*, 390–394. Available online: https://pubmed.ncbi.nlm.nih.gov/18455366/ (accessed on 10 November 2021). [CrossRef]
62. Winkler, I.; Sobol, E.; Yagen, B.; Steinman, A.; Devor, M.; Bialer, M. Efficacy of antiepileptic tetramethylcyclopropyl analogues of valproic acid amides in a rat model of neuropathic pain. *Neuropharmacology* **2005**, *49*, 1110–1120. Available online: https://pubmed.ncbi.nlm.nih.gov/16055160/ (accessed on 18 November 2021). [CrossRef]
63. Blotnik, S.; Bergman, F.; Bialer, M. Disposition of two tetramethylcyclopropane analogues of valpromide in the brain, liver, plasma and urine of rats. *Eur. J. Pharm. Sci.* **1998**, *6*, 93–98. Available online: https://pubmed.ncbi.nlm.nih.gov/9795021/ (accessed on 18 November 2021). [CrossRef]
64. Okada, A.; Onishi, Y.; Yagen, B.; Shimshoni, J.A.; Kaufmann, D.; Bialer, M.; Fujiwara, M. Tetramethylcyclopropyl Analogue of the Leading Antiepileptic Drug, Valproic Acid: Evaluation of the Teratogenic Effects of Its Amide Derivatives in NMRI Mice. *Birth Defects Res. Part A Clin. Mol. Teratol.* **2008**, *82*, 610–621. Available online: https://pubmed.ncbi.nlm.nih.gov/18671279/ (accessed on 18 November 2021). [CrossRef]
65. Pessah, N.; Bialer, M.; Wlodarczyk, B.; Finnell, R.H.; Yagen, B. Alpha-fluoro-2,2,3,3-tetramethylcyclopropanecarboxamide, a Novel Potent Anticonvulsant Derivative of a Cyclic Analogue of Valproic Acid. *J. Med. Chem.* **2009**, *52*, 2233–2242. Available online: https://pubmed.ncbi.nlm.nih.gov/19296679/ (accessed on 18 November 2021). [CrossRef]
66. Shekh-Ahmad, T.; Mawasi, H.; McDonough, J.H.; Yagen, B.; Bialer, M. The Potential of Sec-Butylpropylacetamide (SPD) and Valnoctamide and their Individual Stereoisomers in Status Epilepticus. *Epilepsy Behav.* **2015**, *49*, 298–302. Available online: https://pubmed.ncbi.nlm.nih.gov/25979572/ (accessed on 22 November 2021). [CrossRef]
67. Safety and Efficacy of NPS 1776 in the Acute Treatment of Migraine Headaches-Full Text View-ClinicalTrials.gov. Available online: https://clinicaltrials.gov/ct2/show/NCT00172094?id=NCT00172094&draw=2&rank=1&load=cart (accessed on 28 September 2020).
68. Löscher, W.; Nau, H. Pharmacological Evaluation of Various Metabolites and Analogues of Valproic Acid. Anticonvulsant and Toxic Potencies in Mice. *Neuropharmacology* **1985**, *24*, 427–435. Available online: https://pubmed.ncbi.nlm.nih.gov/3927183/ (accessed on 22 November 2021). [CrossRef]
69. Okada, A.; Aoki, Y.; Kushima, K.; Kurihara, H.; Bialer, M.; Fujiwara, M. Polycomb Homologs are Involved in Teratogenicity of Valproic Acid in Mice. *Birth Defects Res. Part A Clin. Mol. Teratol.* **2004**, *70*, 870–879. Available online: https://pubmed.ncbi.nlm.nih.gov/15523661/ (accessed on 10 September 2020). [CrossRef]
70. Paradis, F.H.; Hales, B.F. Exposure to Valproic Acid Inhibits Chondrogenesis and Osteogenesis in Mid-Organogenesis Mouse Limbs. *Toxicol. Sci.* **2013**, *131*, 234–241. Available online: https://pubmed.ncbi.nlm.nih.gov/23042728/ (accessed on 10 September 2020). [CrossRef]
71. Bialer, M.; Rubinstein, A. Pharmacokinetics of Valpromide in Dogs after Various Modes of Administration. *Biopharm. Drug Dispos.* **1984**, *5*, 177–183. Available online: https://pubmed.ncbi.nlm.nih.gov/6430363/ (accessed on 25 November 2021). [CrossRef]
72. Bialer, M.; Rubinstein, A.; Raz, I.; Abramsky, O. Pharmacokinetics of Valpromide after Oral Administration of a Solution and a Tablet to Healthy Volunteers. *Eur. J. Clin. Pharmacol.* **1984**, *27*, 501–503. Available online: https://pubmed.ncbi.nlm.nih.gov/6440792/ (accessed on 25 November 2021). [CrossRef]
73. Lindekens, H.; Smolders, I.; Khan, G.M.; Bialer, M.; Ebinger, G.; Michotte, Y. In Vivo Study of the Effect of Valpromide and Valnoctamide in the Pilocarpine Rat Model of Focal Epilepsy. *Pharm. Res.* **2000**, *17*, 1408–1413. Available online: https://pubmed.ncbi.nlm.nih.gov/11205735/ (accessed on 10 September 2020). [CrossRef]
74. Kaufmann, D.; Yagen, B.; Minert, A.; Wlodarczyk, B.; Finnell, R.H.; Schurig, V.; Devor, M.; Bialer, M. Evaluation of the Antiallodynic, Teratogenic and Pharmacokinetic Profile of Stereoisomers of Valnoctamide, an Amide Derivative of a Chiral Isomer of Valproic Acid. *Neuropharmacology* **2010**, *58*, 1228–1236. Available online: https://pubmed.ncbi.nlm.nih.gov/20230843/ (accessed on 10 September 2020). [CrossRef]
75. Okada, A.; Kushima, K.; Aoki, Y.; Bialer, M.; Fujiwara, M. Identification of Early-Responsive Genes Correlated to Valproic Acid-Induced Neural Tube Defects in Mice. *Birth Defects Res. Part A Clin. Mol. Teratol.* **2005**, *73*, 229–238. Available online: https://pubmed.ncbi.nlm.nih.gov/15799026/ (accessed on 10 September 2020). [CrossRef]
76. Bialer, M.; White, H.S. Key Factors in the Discovery and Development of New Antiepileptic Drugs. *Nat. Rev. Drug Discov.* **2010**, *9*, 68–82. Available online: https://pubmed.ncbi.nlm.nih.gov/20043029/ (accessed on 25 November 2021). [CrossRef]
77. Bialer, M.; Yagen, B. Valproic Acid: Second Generation. *Neurotherapeutics* **2007**, *4*, 130–137. Available online: https://pubmed.ncbi.nlm.nih.gov/17199028/ (accessed on 25 November 2021). [CrossRef]
78. Kaufmann, D.; West, P.J.; Smith, M.D.; Yagen, B.; Bialer, M.; Devor, M.; White, H.S.; Brennan, K.C. sec-Butylpropylacetamide (SPD), a New Amide Derivative of Valproic Acid for the Treatment of Neuropathic and Inflammatory Pain. *Pharmacol. Res.* **2017**, *117*, 129–139. Available online: https://pubmed.ncbi.nlm.nih.gov/27890817/ (accessed on 25 November 2021). [CrossRef]
79. Hen, N.; Shekh-Ahmad, T.; Yagen, B.; McDonough, J.H.; Finnell, R.H.; Wlodarczyk, B.; Bialer, M. Stereoselective Pharmacodynamic and Pharmacokinetic Analysis of sec-Butylpropylacetamide (SPD), a New CNS-Active Derivative of Valproic Acid with Unique Activity against Status Epilepticus. *J. Med. Chem.* **2013**, *2256*, 6467–6477. Available online: https://pubmed.ncbi.nlm.nih.gov/23879329/ (accessed on 22 November 2021). [CrossRef]

80. Haj-Yehia, A.; Bialer, M. Structure–Pharmacokinetic Relationships in a Series of Valpromide Derivatives with Antiepileptic Activity. *Pharm. Sci.* **1989**, *6*, 683–689. Available online: https://pubmed.ncbi.nlm.nih.gov/2510141/ (accessed on 10 September 2020).
81. Isoherranen, N.; Yagen, B.; Woodhead, J.H.; Spiegelstein, O.; Blotnik, S.; Wilcox, K.S.; Finnell, R.H.; Bennett, G.D.; White, H.S.; Bialer, M. Characterization of the Anticonvulsant Profile and Enantioselective Pharmacokinetics of the Chiral Valproylamide Propylisopropyl Acetamide in Rodents. *Br. J. Pharmacol.* **2003**, *138*, 602–613. Available online: https://pubmed.ncbi.nlm.nih.gov/12598414/ (accessed on 10 September 2020). [CrossRef]
82. Mareš, P.; Kubová, H.; Hen, N.; Yagen, B.; Bialer, M. Derivatives of Valproic Acid are Active against Pentetrazol-Induced Seizures in Immature Rats. *Epilepsy Res.* **2013**, *106*, 64–73. Available online: https://pubmed.ncbi.nlm.nih.gov/23815889/ (accessed on 10 September 2020). [CrossRef]
83. Spiegelstein, O.; Yagen, B.; Levy, R.H.; Finnell, R.H.; Bennett, G.D.; Roeder, M.; Schurig, V.; Bialer, M. Stereoselective Pharmacokinetics and Pharmacodynamics of Propylisopropyl Acetamide, a CNS-Active Chiral Amide Analog of Valproic Acid. *Pharm. Res.* **1999**, *16*, 1582–1588. Available online: https://pubmed.ncbi.nlm.nih.gov/10554101/ (accessed on 10 September 2020). [CrossRef]
84. Spiegelstein, O.; Bialer, M.; Radatz, M.; Nau, H.; Yagen, B. Enantioselective Synthesis and Teratogenicity of Propylisopropyl Acetamide, a CNS-Active Chiral Amide Analogue of Valproic Acid. *Chirality* **1999**, *11*, 645–650. Available online: https://pubmed.ncbi.nlm.nih.gov/10467316/ (accessed on 10 September 2020). [CrossRef]
85. Bialer, M.; Johannessen, S.I.; Levy, R.H.; Perucca, E.; Tomson, T.; White, H.S. Progress Report on New Antiepileptic Drugs: A Summary of the Tenth Eilat Conference (EILAT X). *Epilepsy Res.* **2010**, *92*, 89–124. Available online: http://www.ncbi.nlm.nih.gov/pubmed/20970964. (accessed on 22 November 2021). [CrossRef]
86. Loescher, W. Anticonvulsant Activity of Metabolites of Valproic Acid-PubMed. Available online: https://pubmed.ncbi.nlm.nih.gov/6784685/ (accessed on 22 November 2021).
87. Loscher, W.; Nau, H. Distribution of Valproic Acid and Its Metabolites in Various Brain Areas of Dogs and Rats after Acute and Prolonged Treatment-PubMed. Available online: https://pubmed.ncbi.nlm.nih.gov/6411902/ (accessed on 22 November 2021).
88. Löscher, W.; Nau, H.; Marescaux, C.; Vergnes, M. Comparative Evaluation of Anticonvulsant and Toxic Potencies of Valproic Acid and 2-en-valproic Acid in Different Animal Models of Epilepsy. *Eur. J. Pharmacol.* **1984**, *99*, 211–218. Available online: https://pubmed.ncbi.nlm.nih.gov/6428923/ (accessed on 10 September 2020). [CrossRef]
89. Strichartz, G.R.; Wang, G.K. Rapid Voltage-Dependent Dissociation of Scorpion Alpha-Toxins Coupled to Na Channel Inactivation in Amphibian Myelinated Nerves. *J. Gen. Physiol.* **1986**, *88*, 413–435. Available online: https://pubmed.ncbi.nlm.nih.gov/2428923/ (accessed on 22 November 2021). [CrossRef]
90. Isoherranen, N.; Yagen, B.; Bialer, M. New CNS-Active Drugs Which are Second-Generation Valproic Acid: Can They Lead to the Development of a Magic Bullet? *Curr. Opin. Neurol.* **2003**, *16*, 203–211. Available online: https://pubmed.ncbi.nlm.nih.gov/12644750/ (accessed on 10 September 2020). [CrossRef]
91. Tang, W.; Abbott, F.S. A Comparative Investigation of 2-Propyl-4-Pentenoic Acid (4-ene VPA) and its α-Fluorinated Analogue: Phase II Metabolism and Pharmacokinetics. *Drug Metab. Dispos.* **1997**, *25*, 219–227.
92. Kesterson, J.W.; Granneman, G.R.; Machinist, J.M. The Hepatotoxicity of Valproic acid and Its Metabolites in Rats. I. Toxicologic, Biochemical and Histopathologic Studies. *Hepatology* **1984**, *4*, 1143–1152. [CrossRef]
93. Sobol, E.; Bialer, M.; Yagen, B. Tetramethylcyclopropyl Analogue of a Leading Antiepileptic Drug, Valproic acid. Synthesis and Evaluation of Anticonvulsant Activity of Its Amide Derivatives. *J. Med. Chem.* **2004**, *47*, 4316–4326. Available online: https://pubmed.ncbi.nlm.nih.gov/15294003/ (accessed on 18 November 2021). [CrossRef]
94. Bialer, M.; Hadad, S.; Kadry, B.; Abdul-Hai, A.; Haj-Yehia, A.; Sterling, J.; Herzig, Y.; Yagen, B. Pharmacokinetic Analysis and Antiepileptic Activity of Tetra-Methylcyclopropane Analogues of Valpromide. *Pharm. Res.* **1996**, *13*, 284–289. Available online: https://pubmed.ncbi.nlm.nih.gov/8932450/ (accessed on 18 November 2021). [CrossRef]
95. Bialer, M.; Johannessen, S.I.; Levy, R.H.; Perucca, E.; Tomson, T.; White, H.S. Progress Report on New Antiepileptic Drugs: A Summary of the Ninth Eilat Conference (EILAT IX). *Epilepsy Res.* **2009**, *83*, 1–43. Available online: https://pubmed.ncbi.nlm.nih.gov/19008076/ (accessed on 25 November 2021). [CrossRef]
96. Blotnik, S.; Bergman, F.; Bialer, M. The Disposition of Valproyl Glycinamide and Valproyl Glycine in Rats. *Pharm. Res.* **1997**, *14*, 873–878. [CrossRef]
97. Bialer, M.; Johannessen, S.I.; Kupferberg, H.J.; Levy, R.H.; Loiseau, P.; Perucca, E. Progress Report on New Antiepileptic Drugs: A Summary of the Fifth Eilat Conference (EILAT V). *Epilepsy Res.* **2001**, *43*, 11–58. Available online: https://pubmed.ncbi.nlm.nih.gov/11137386/ (accessed on 22 November 2021). [CrossRef]
98. Wlodarczyk, B.J.; Ogle, K.; Lin, L.Y.; Bialer, M.; Finnell, R.H. Comparative Teratogenicity Analysis of Valnoctamide, Risperidone, and Olanzapine in Mice. *Bipolar Disord.* **2015**, *17*, 615–625. Available online: https://pubmed.ncbi.nlm.nih.gov/26292082/ (accessed on 26 November 2021). [CrossRef]
99. Bialer, M.; Johannessen, S.I.; Levy, R.H.; Perucca, E.; Tomson, T.; White, H.S. Progress Report on New Antiepileptic Drugs: A Summary of the Eleventh Eilat Conference (EILAT XI). *Epilepsy Res.* **2013**, *103*, 2–30. Available online: https://pubmed.ncbi.nlm.nih.gov/23219031/ (accessed on 26 November 2021). [CrossRef]

100. Mawasi, H.; Shekh-Ahmad, T.; Finnell, R.H.; Wlodarczyk, B.J.; Bialer, M. Pharmacodynamic and Pharmacokinetic Analysis of CNS-Active Constitutional Isomers of Valnoctamide and sec-Butylpropylacetamide—Amide Derivatives of Valproic Acid. *Epilepsy Behav.* **2015**, *46*, 72–78. Available online: https://pubmed.ncbi.nlm.nih.gov/25863940/ (accessed on 26 November 2021). [CrossRef]
101. Pouliot, W.; Bialer, M.; Hen, N.; Shekh-Ahmad, T.; Kaufmann, D.; Yagen, B.; Ricks, K.; Roach, B.; Nelson, C.; Dudek, F.E. A Comparative Electrographic Analysis of the Effect of sec-butyl-propylacetamide on Pharmacoresistant Status Epilepticus. *Neuroscience* **2013**, *231*, 145–156. Available online: https://pubmed.ncbi.nlm.nih.gov/23159312/ (accessed on 26 November 2021). [CrossRef]

J Mater Sci: Mater Electron

Investigation of charge transport mechanism in hydrothermally synthesized reduced graphene oxide (rGO) incorporated zinc oxide (ZnO) nanocomposite films

Nikita Jain[1], Savita Sharma[2], and Nitin K. Puri[1,*]

[1] *Nanomaterials Research Laboratory (NRL), Department of Applied Physics, Delhi Technological University, Delhi 110042, India*
[2] *Physics Department, Kalindi College, University of Delhi, Delhi 110008, India*

Received: 26 July 2021
Accepted: 19 November 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

ABSTRACT

The present study aims to investigate the impact of reduced graphene oxide (rGO) incorporation on the charge transport properties of zinc oxide (ZnO) nanocomposite films. ZnO and varied weight percentage of rGO (1.25% to 10%) in ZnO-rGO nanocomposites are synthesized via cost-effective and facile hydrothermal method. The effect of varying weight percentage of rGO in ZnO nanocomposite is analysed by techniques such as X-ray diffraction (XRD), Scanning electron microscopy (SEM), Energy dispersive X-ray (EDX), Fourier transform infra-red spectroscopy (FTIR), and Raman spectroscopy. The observed current–voltage (I–V) characteristics at room temperature show the enhancement in forward current with an increasing weight percentage of rGO (1.25% to 10%) in ZnO nanocomposite films. To study the charge transport mechanism in nanocomposite films, dual-logarithmic I–V characteristics are plotted. From the characteristic curves, we find that three different laws of space charge limited conduction (SCLC) model namely Ohm's law, Child's law, and trap-limited SCLC mechanism describe charge transport properties in the ZnO-rGO nanocomposite films. At a low weight percentage of rGO (1.25%) in ZnO-rGO nanocomposite films, a transition from Child's law to trap-limited SCLC mechanism (0.9 V being the cross-over voltage) is obtained. As the weight percentage of rGO in ZnO-rGO nanocomposite films is increased from 2.5 to 10%, the conduction is favored by Ohm's law at low applied voltages to Child's law at higher applied voltages. Best experimental results are shown with 5% of rGO in ZnO-rGO nanocomposite. The prepared nanocomposite films have potential applications in UV-photodetector devices.

Address correspondence to E-mail: nitinkumarpuri@dtu.ac.in

https://doi.org/10.1007/s10854-021-07445-6
Published online: 02 December 2021

1 Introduction

Nanocomposites are a blend of dissimilar materials with one or more of their constituents being at the nanoscale range. They offer an opportunity to tailor the properties of material towards a specific application. There has been an enormous interest in nanocomposite materials in recent times, owing to their cost-effectiveness and enhanced properties leading to improved device performance. Hybrid nanocomposites have been developed and explored for profound applications such as photodetectors, supercapacitors, gas sensors, solar cells. [1, 2]. Among the various metal-oxides, zinc oxide (ZnO) is a prominent semi-conductor with a wide range of applications due to its wide-bandgap, large exciton binding energy, abundant availability, non-toxicity, and high photon-absorption [3, 4]. A variety of morphologies can be synthesized by numerous chemical and physical methods widening up the research opportunities towards new technology [4]. Simultaneously, easy tuning of shape, size, and compositions of ZnO play an important role in altering its magnetic, optical, electrical, and other properties for numerous applications [5, 6].

Reduced graphene oxide (rGO), a two-dimensional (2D) material is an expeditious rising star on the horizon of material science and nanotechnology. Its exceptional mobility and conductivity, mechanical flexibility, optical transparency, chemical stability, and high specific surface area is attracting considerable research interest [7]. The sheets of rGO possess a large surface area which may provide support to load other functional nanomaterials. Furthermore, the oxygenated functional groups can act as nucleation centers to support nanomaterials onto graphene sheets which makes it viable to synthesize nanocomposites based on graphene [8].

Various hybrid nanocomposites based on ZnO and rGO have been explored in the field of photodetectors, gas sensors, batteries, supercapacitors, and many more due to their enhanced properties [9–14]. Previously, it has been widely reported about the charge transport properties of ZnO and rGO materials individually and their nanocomposites with other different materials (eg. n-type ZnO grown on plain and focussed ion beam GaN substrates, n-Gallium doped-ZnO on p-Silicon substrate, GaN/rGO nanocomposite with a different weight percentage of GaN, etc. [15–17]). Soylu et al. studied the effect of the molar concentration of precursors of ZnO on I–V characteristics [18]. Therefore, it is evident that ZnO, rGO, and their nanocomposites with other materials have been explored extensively. However, to the best of our knowledge, no reports are available focussing on a detailed study of the charge transport mechanism in ZnO-rGO nanocomposite films.

A plethora of experimental methods have been rigorously explored for binding of ZnO with rGO such as hydrothermal, aerosol spraying, microwave synthesis, chemical deposition, solvothermal, precipitation methods, and solution combustion synthesis [19–23]. Among these methods, hydrothermal synthesis is considered advantageous as it is cost-effective, uses mild synthesis conditions, and requires low energy with simple equipments resulting in uniform ZnO nanoparticles over rGO nanosheets [24].

In the present work, ZnO and ZnO-rGO nanocomposites with different weight percentage of rGO (from 1.25 to 10%) have been synthesized using hydrothermal method. The charge transport mechanism in the prepared Al/ZnO/ITO and Al/ZnO-rGO/ITO nanocomposite films (rGO varied from 1.25 to 10%) have been studied in detail towards their utilization for UV-photodetector devices.

2 Experimental

2.1 Chemical reagents

The chemicals used in the synthesis are Zinc acetate ($Zn(CH_3COO)_2{\cdot}2H_2O$), sodium hydroxide pellets (NaOH), graphite powder, concentrated sulphuric acid (H_2SO_4), orthophosphoric acid (H_3PO_4), hydrogen peroxide (H_2O_2), potassium permanganate ($KMnO_4$), hydrazine hydrate ($N_2H_4{\cdot}H_2O$), and ethanol (C_2H_5OH). All the chemicals are purchased from Fisher Scientific and Sigma Aldrich. All belong to AR grade purity and are employed without further purification.

2.2 Synthesis of ZnO twin-hexapods

Firstly, ZnO twin-hexapods are synthesized using the facile hydrothermal method. In the synthesis, a solution of Zinc acetate (0.2 M) and sodium hydroxide (1.25 M) is prepared separately, with continuous

magnetic stirring for 30 min. Then, the solution of sodium hydroxide is added into zinc acetate solution drop-wise. The pH of the solution is observed to be 6.6. The final solution is kept for stirring for an hour which is then transferred to a 100 ml Teflon-lined steel autoclave, which is heated at 130 °C for 24 h. After several washings with DI and ethanol, the final product is dried overnight.

2.3 Synthesis of rGO

Firstly, GO is synthesized from graphite using improved modified Hummer's method as reported elsewhere [25, 26]. In detail, 2 g of graphite powder is added to a mixture of concentrated sulphuric acid and phosphoric acid in the ratio of 9:1 with continuous stirring. Thereafter, some amount of potassium permanganate is added at a very slow pace to the resulting solution followed by the magnetic stirring of 12 h at 50 °C. After this, the reaction is quenched by adding ice (300 ml) with 2.5 ml of hydrogen peroxide solution. Finally, the yellowish slurry mixture is washed (until the pH reached to 7), filtered, and dried to obtain the final product. For the synthesis of rGO, 200 mg of GO powder so obtained is dispersed in 200 ml of DI water and ultrasonicated for an hour. Then it is transferred onto the stirrer and 100 µL of hydrazine hydrate is added drop-wise into it which acts as a reducing agent and the solution is left for 12 h stirring at a particular temperature. The final solution turns black which then is washed, filtered, and dried to obtain rGO.

2.4 Synthesis of ZnO-rGO nanocomposites

Various nanocomposites of ZnO-rGO are synthesized via hydrothermal route by varying weight percentage of rGO (from 1.25 to 10%) in ZnO. Typically, a suitable amount of rGO (1 mg/ml) is dispersed in ethanol with ultrasonic treatment for an hour. Meanwhile, a solution of Zinc acetate (0.2 M) and sodium hydroxide (1.25 M) is prepared separately with stirring. It is followed by mixing of all three resulting solutions under continuous stirring for an hour. The observed pH of the solution (ZnO-rGO nanocomposite) is 6.6. At last, the solution is transferred to a Teflon-lined stainless-steel autoclave which is heated at 130 °C for 24 h. After several washings with DI water and ethanol, the grey color powder is obtained which is then dried overnight. Weight percent variation by volume (w/v) of 1.25%, 2.5%, 5%, and 10% of rGO to ZnO are prepared and named as ZG-1.25, ZG-2.5, ZG-5, and ZG-10, respectively. All these ZnO-rGO nanocomposites with various rGO content are prepared by above said process.

2.5 Film fabrication

ZnO-rGO nanocomposite films are fabricated on rectangular indium tin-oxide (ITO) coated corning-glass substrates. The standard wet cleaning method is employed to clean the ITO substrates. Briefly, soap solution, deionized ionized (DI) water, acetone, and finally with isopropyl alcohol. After cleaning, the substrates are dried at 100 °C on a hot plate for 30 min. In the first step, ZnO powder is dispersed in ethanol by ultrasonicating it for 30 min. In the second step, ZnO films are prepared via spin-coating technique, by dispensing the above solution on ITO substrates at 3000 rpm for 60 s. It is followed by drying the films in the open air at 70 °C for 10 min. A similar process is devised to obtain all the desired ZG-1.25, ZG-2.5, ZG-5, and ZG-10 nanocomposites films. Finally, 100 nm thick top Al metal electrodes are deposited on the prepared films using the thermal evaporation technique. Figure 1 shows the schematic description of the process of fabrication of Al/ZnO-rGO/ITO nanocomposites films.

3 Results and discussions

All the prepared samples are thoroughly characterized using various techniques such as X-ray diffraction (XRD), Scanning electron microscopy (SEM), Energy dispersive x-ray (EDX), Fourier transform infra-red spectroscopy (FTIR), and Raman spectroscopy. Charge transport properties of all the prepared films (Al/ZnO/ITO, Al/ZnO-rGO/ITO) have been studied via current–voltage (I–V) characteristics over the voltage range of − 2.5 to 2.5 V using Keithley 2450 SMU.

3.1 Structural analysis (X-ray diffraction)

The XRD pattern of ZnO, ZG-1.25, ZG-2.5, ZG-5, and ZG-10 nanocomposite is shown in Fig. 2. XRD pattern of ZnO shows all the diffraction peaks which are congruous with the standard data available for the

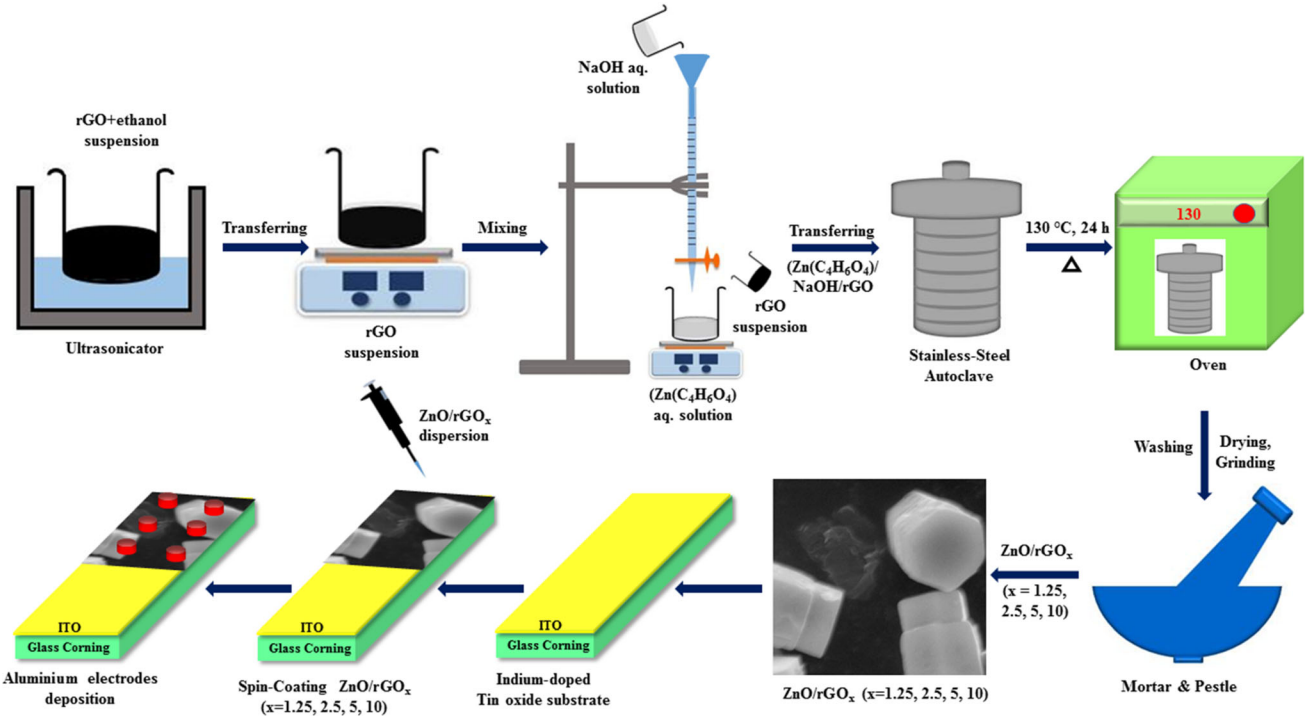


Fig. 1 Schematic description of fabrication of Al/ZnO-rGO/ITO nanocomposites films

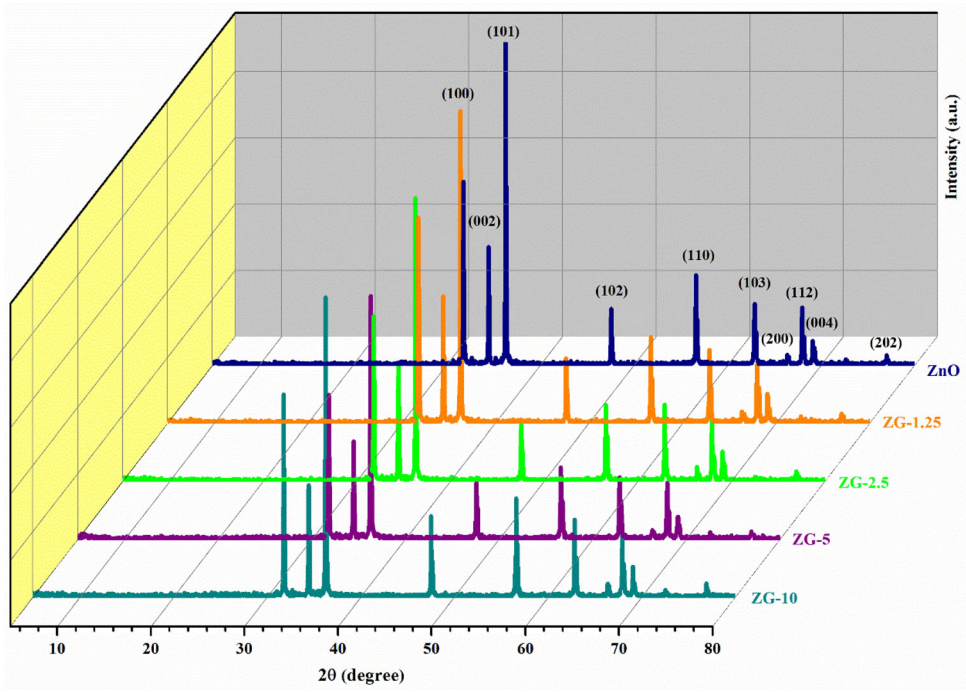


Fig. 2 XRD pattern of ZnO, ZG-1.25, ZG-2.5, ZG-5, and ZG-10 nanocomposites

wurtzite structure of ZnO (JCPDS 36-1451) [27]. A prominent diffraction peak of GO (Fig. S1a in the Supplementary Information) is found to be at 11.7° and a weak peak located at 42.6° corresponding to (002) and (100) reflection, respectively. The inset of Fig. S1a in the Supplementary Information shows the diffraction pattern of raw graphite which shows a highly crystalline and strong diffraction peak at 26.6° towards the (002) plane. The shift in GO peak at a lower diffraction angle confirms the successful

oxidation of graphite [28]. The broadening and shifting of the distinct peak of GO from 11.7° to 23.1° can be attributed to the breakdown of the long-range order of GO [29]. Two broad diffraction peaks located at 23.1° and 43.6° correspond to rGO (Fig. S1b in the Supplementary Information). These two peaks confirm the formation of (002) and (100) planes of rGO [29]. The peak positions of ZnO remain the same in all ZnO-rGO nanocomposites (Fig. 2) while the two reflection peaks of rGO are absent in nanocomposites. The reason can be attributed to the low diffraction intensity of rGO in the dispersed state [30, 31].

The d-spacing of graphite, GO and rGO is calculated using Bragg's relation [32] and is found to be 0.33 nm, 0.75 nm, and 0.38 nm, respectively for the (002) plane. The increase of d-spacing of GO compared to graphite can be due to the intercalation of oxygenated functional groups (such as epoxide and hydroxyl groups) and intercalation of molecules of water in the interlayer spacing of carbon layers [28].

The average crystallite size of ZnO and ZnO-rGO nanocomposites is calculated using Scherrer's relation [33, 34]. The crystallite size calculated for ZnO, ZG-1.25, ZG-2.5, ZG-5, and ZG-10 samples is 81, 77, 76, 70, and 74 nm, respectively. It is observed that the crystallite size of ZnO decreases progressively as the weight percentage of rGO in the nanocomposites is increased to 5%. This suggests that the growth and nucleation stage of ZnO crystals in the nanocomposites is hindered by rGO [4]. Furthermore, the dislocation density (δ) of ZnO and ZnO-rGO nanocomposite is calculated using the following relation: $\delta = 1/D^2$ where D is the crystallite size of ZnO and ZnO-rGO nanocomposite. The highest and lowest dislocation density is found to be approximately 2.05×10^{-4} nm^{-2} and 1.49×10^{-4} nm^{-2} for ZG-5 nanocomposite and ZnO, respectively. When the dislocation density decreases, the crystallinity increases [35]. The intensity of the diffraction peak along the (101) diffraction plane is supporting this phenomenon. The diffraction peaks of the composite ZG-1.25 are stronger than those of other nanocomposites but weaker than the diffraction peak of pure ZnO which indicates that the crystallinity of the composite is affected by the addition of rGO [20].

The lattice parameters (lattice constants and volume of unit cell) are calculated by means of Rietveld refinement using FullProf Software. Best matching of the peaks for all the samples is observed for hexagonal structure with $P6_3mc$ space group. Table 1 shows the calculated values of FWHM, lattice parameters, crystallite size, and dislocation density for all the prepared samples. It is clear from the table that the structural lattice parameters increases and the crystallite size decreases with the increasing weight percentage of rGO in the ZnO-rGO nanocomposites upto 5%. The obtained results are consistent with the literature [36–38].

On the other hand, the variation in the trend of certain crystal parameters of ZG-10 nanocomposite can be understood as: at this concentration, the ratio of zinc ions to rGO is not critical. Since the critical ratio is not maintained, the sorption of zinc ions is less on rGO nanosheets, due to which the synergetic effect could not follow the trend as followed by the other nanocomposites.

3.2 SEM and EDX observations

SEM analysis is done to investigate the surface morphology as well as to examine the average length of the prepared twin-hexapods of ZnO, ZG-1.25, ZG-2.5, ZG-5, and ZG-10 nanocomposites. Figure 3 shows the SEM images of the prepared ZnO and ZnO-rGO nanocomposite with variation in weight percentage of rGO from 1.25 to 10%. It is observed that the average length of the hexapods is closely related to the weight percentage of rGO in the ZnO-rGO nanocomposite. Figure 3a depicts the SEM image of ZnO. The observed average length of the top-hexapod is 2.7 μm and the average length of the bottom-hexapod is 5 μm. Compared to the crystallite size of the samples calculated by the XRD analysis, a bigger particle size of ZnO and ZnO-rGO nanocomposites (rGO varied from 1.25 to 10%) is obtained from SEM observations. This can be due to the fact that particles contain many smaller crystallites [39, 40]. When rGO is introduced in the ZnO framework, the morphology is maintained as can be seen (Fig. 3b–e). The top-hexapod of the ZG-1.25 nanocomposite (Fig. 3b) has an average length of 1.7 μm and the average length of the bottom-hexapod is 3.1 μm. Thus, we can clearly see that the size of ZnO hexapods decreases with the increase in the concentration of rGO in nanocomposites. Top-hexapods of average length 1.6 μm and bottom-hexapod of the average length of 2.9 μm is obtained when the rGO concentration is increased to 2.5% in ZnO (Fig. 3c). But, further increase in rGO concentration to 5% in ZnO-rGO nanocomposite the ZnO twin-hexapods are found to be comparatively

Table 1 Variation in FWHM, lattice parameters (lattice constants and volume of unit cell), crystallite size and dislocation density of (a) ZnO, (b) ZG-1.25, (c) ZG-2.5, (d) ZG-5, and (e) ZG-10 nanocomposites with variation in weight percentage of rGO

Sample	Plane	2θ (degree)	FWHM (degree)	Lattice parameters		Volume of unit cell (Å^3)	Crystallite size (nm)	Dislocation density (nm^{-2}) $\times 10^{-4}$
				a (Å)	c (Å)			
ZnO	(101)	36.3	0.10	3.2465	5.2021	47.48	81	1.49
ZG-1.25	(101)	36.3	0.10	3.2477	5.2034	47.53	77	1.67
ZG-2.5	(101)	36.3	0.11	3.2481	5.2040	47.54	76	1.74
ZG-5	(101)	36.3	0.11	3.2485	5.2041	47.55	70	2.05
ZG-10	(101)	36.3	0.11	3.2480	5.2039	47.54	74	1.82

small with an average length of 1.2 μm and 2.4 μm (Fig. 3d) of top and bottom-hexapod, respectively. The reason is the formation of rGO sheets which creates a hindrance for further growth of ZnO hexapods which is consistent with the XRD analysis as well.

On the other hand, when the rGO concentration is increased to 10% in ZnO-rGO nanocomposite, most of the rGO sheets in the nanocomposite are agglomerated and some of the ZnO grow on graphene sheets. To understand and deduce the ongoing observation in surface morphology, the formation process of ZnO-rGO nanocomposite is surmised. There should be a critical ratio of zinc ions of ZnO and rGO to form well-dispersed nanocomposites. When the concentration of rGO is low, the ratio of zinc ions is relatively high in rGO. This leads to excess sorption of zinc ions on rGO sheets which results in well-dispersed ZnO-rGO nanocomposite. On increasing the concentration of rGO in ZnO, the sorption of zinc ions is less on rGO sheets. The negative charges on reduced graphene oxide are partially neutralized by the zinc ions. This lead to the coagulation of rGO sheets during the reaction process [41]. It is observed that the top-hexapod of ZnO in this nanocomposite has an average length of 1.5 μm and the bottom-hexapod has an average length of 2.7 μm (Fig. 3e). But, an increase in rGO concentration in the nanocomposites does not result in the agglomeration of ZnO hexapods. From the above observation, we can deduce that rGO nanosheets co-exist with the ZnO twin-hexapods in the fabricated ZnO-rGO nanocomposites.

The elemental composition of ZnO and ZnO-rGO nanocomposites are analysed by the EDX spectrum. It is inapt to determine the exact content of the elements present in the ZnO and ZnO-rGO nanocomposites using EDX because of the presence of light elements (O and C) [42]. EDX spectra and the weight percent composition of ZnO and ZnO-rGO nanocomposites are shown in Fig. S2 in Supplementary Information. As can be seen from the EDX spectra of ZnO (S2a in Supplementary Information), no peak corresponding to carbon is found. With the addition of rGO in the ZnO framework, increasing weight percent of carbon peak is observed in all the ZnO-rGO nanocomposites (Fig. S2b–e in Supplementary Information). However, an impurity peak of sodium is also found in all the samples. This impurity peak could be due to the precursor (NaOH) used in the synthesis of ZnO and ZnO-rGO nanocomposites. Although its weight percent composition is negligibly small, yet it may be one of the reasons for the creation of traps inside ZnO and ZnO-rGO nanocomposite [43].

3.3 FTIR Spectroscopy analysis

FTIR Spectroscopy is performed to investigate and elucidate the presence of functional groups related to ZnO, GO, rGO along with the other functional groups which may arise after hybridization in prepared samples. Figure 4 shows the FTIR spectra of ZnO, ZG-1.25, ZG-2.5, ZG-5, and ZG-10 nanocomposites. The spectrum of ZnO (Fig. 4a) exhibits a broad peak at 3278 cm^{-1} which corresponds to O–H stretching vibrations [44]. Few narrow peaks 2830 to 2955 cm^{-1} correspond to C–H bonds [45]. Several peaks appearing between 1360 and 1710 cm^{-1} can be attributed to the stretching modes of COOH group, which are believed to be adsorbed on the surface of

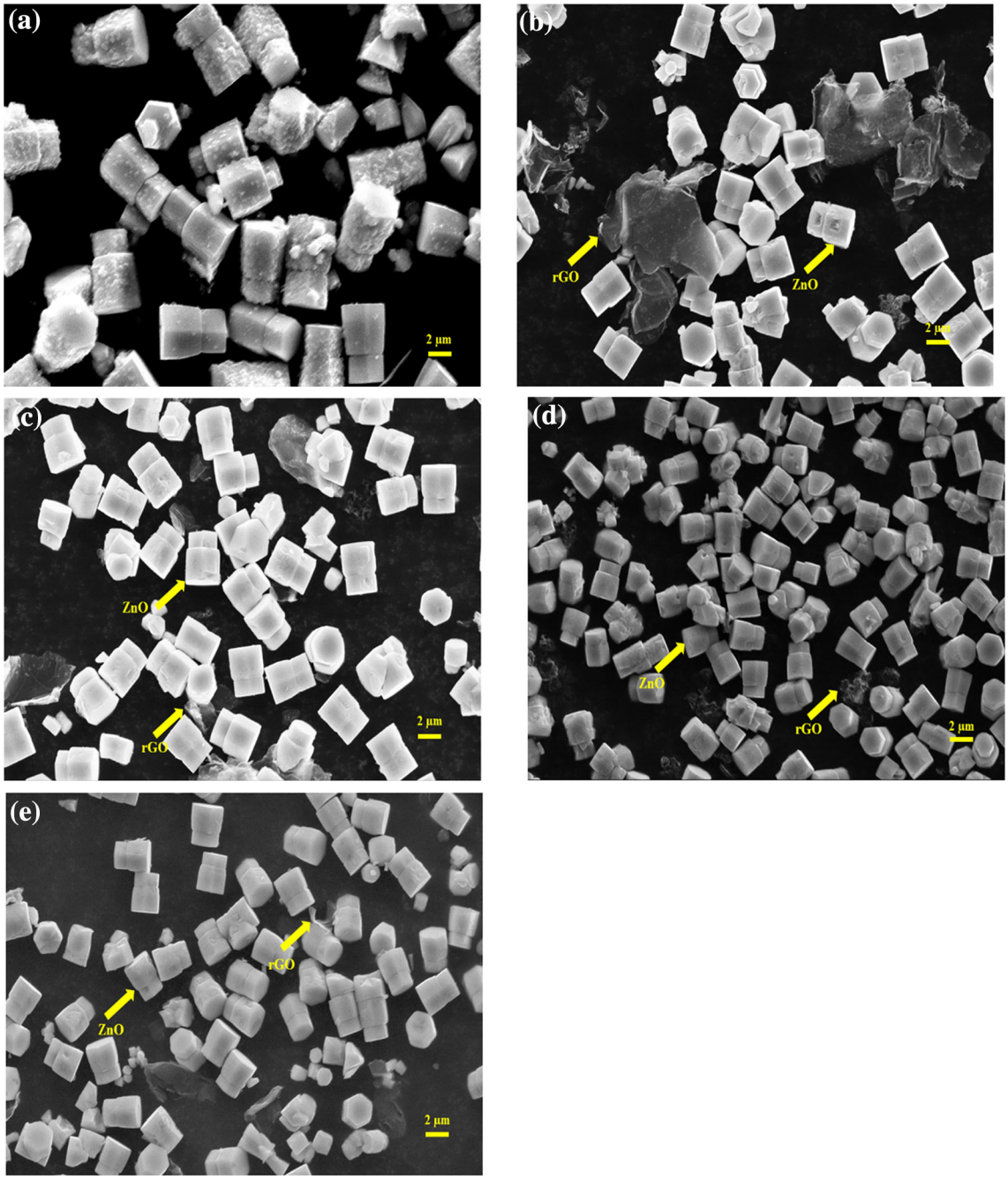


Fig. 3 SEM images of ZnO and ZnO-rGO nanocomposites: **a** ZnO, **b** ZG-1.25, **c** ZG-2.5, **d** ZG-5, and **e** ZG-10

the ZnO hexapods during the process of synthesis [46]. In addition to this, a small peak at 920 cm^{-1} and a pointed peak at 457 cm^{-1} corresponds to Zn–O stretching and deformation vibrations, respectively [46, 47]. FTIR spectrum of GO (Fig. S3a in the Supplementary Information) consists of broad peak 3290 cm^{-1} corresponding to O–H groups and narrow peaks at 2868 cm^{-1} corresponding to C–H stretching. In addition to this, the peak at 1705 cm^{-1} and 1577 cm^{-1} corresponds to C=O and C=C stretching. A small peak at 1395 cm^{-1} corresponds to O–H deformation. Finally, two peaks at 1230 and 1056 cm^{-1} confirm the epoxy C–O–C and alkoxy C–O stretching vibrations, respectively [48]. When GO is reduced to rGO using hydrazine hydrate, the oxygen-containing functional groups are significantly reduced and weakened which is clear in the FTIR spectrum of rGO (Fig. S3b in the Supplementary

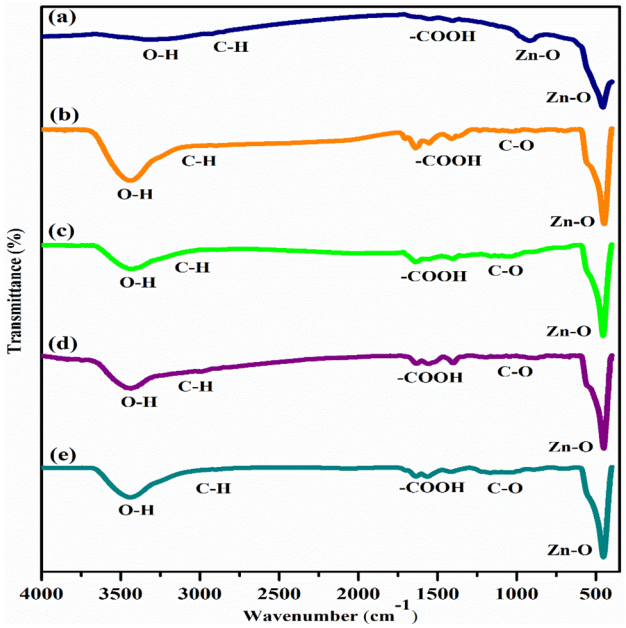


Fig. 4 FTIR spectra of ZnO and ZnO-rGO nanocomposites: a ZnO, b ZG-1.25, c ZG-2.5, d ZG-5, and e ZG-10

Information). The peak at 3450 cm^{-1} is attributed to O–H group. The two new peaks at 1640 cm^{-1} and 1369 cm^{-1} arise due to the cyclic hexagonal symmetry and the vibrations of skeletal of rGO, respectively [47]. Two peaks at 1230 and 1056 cm^{-1} confirm the epoxy C–O–C and alkoxy C–O stretching vibrations, respectively which are also present in GO. The FTIR spectrum of ZnO-rGO nanocomposites (Fig. 4b – e) contains a broad absorption band at 3440 cm^{-1} denoting the O–H groups and a few tiny peaks around 2900 cm^{-1} confirming C-H bonds. Various peaks in the range of 1350 to 1600 cm^{-1} are due to the stretching modes of the COOH group. Furthermore, few weak peaks around 1100 cm^{-1} are attributed to C–O stretching vibrations of rGO sheets. A peak around 457 cm^{-1} is the characteristic peak of ZnO and corresponds to the Zn–O bond confirms the presence of ZnO in all the nanocomposites [9, 21]. Hence, this indicates the coexistence of ZnO and rGO phase in all the prepared ZnO-rGO nanocomposites which is in good accordance with the SEM observations as well.

3.4 Raman Spectroscopy analysis

The Raman spectra of ZnO, GO, rGO and ZG-1.25, ZG-2.5, ZG-5, and ZG-10 nanocomposites is shown in Fig. 5 and the detailed parameters are calculated from Raman data are summarized in Table 2. Raman spectroscopy measurements of ZnO (Fig. 5a) show peaks at 332, 387, 437, 559, and 1145 cm^{-1}. The peak with Raman shift of 332 cm^{-1} corresponds to acoustic and optical phonon overtone with A_1 symmetry while the peak located at 387 cm^{-1} corresponds to A_1(TO) mode [30]. The peak at 437 cm^{-1} corresponds to the ZnO non-polar optical phonons (E_{2high}). It signifies the crystal quality and is the characteristic peak of hexagonal wurtzite phase of ZnO [49]. A peak at 559 cm^{-1} confirms the A_1(LO) mode which is attributed to the surface defect formation of ZnO [50]. A spectral peak at 1145 cm^{-1} is attributed to the multi-phonons process [51].

Two characteristic peaks of GO are observed at 1352 cm^{-1} and 1600 cm^{-1} which corresponds to the D and G bands of graphene, respectively (Fig. S4a in the Supplementary Information). The D band represents lattice disorders in the sp^2-hybridized C atoms and the G band is related to the highly oriented sp^2 hexagonal graphitic lattice [52]. Structural changes from GO to rGO upon chemical reduction are reflected in the spectrum clearly (Fig. S4b in the Supplementary Information). The D and G bands of rGO are located at 1347 cm^{-1} and 1575 cm^{-1}, respectively. The intensity ratio (I_D/I_G) increased from 0.99 to 1.07 when GO is chemically reduced to rGO indicating the partial modification of surface

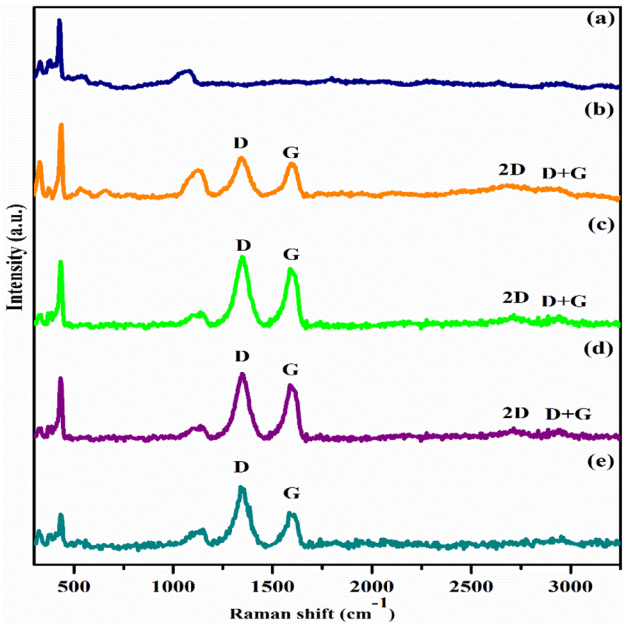


Fig. 5 Raman spectra of ZnO and ZnO-rGO nanocomposites: a ZnO, b ZG-1.25, c ZG-2.5, d ZG-5, and e ZG-10

Table 2 Parameters calculated from Raman data of (a) GO, (b) rGO, (c) ZG-1.25, (d) ZG-2.5, (e) ZG-5, and (f) ZG-10 nanocomposites

Sample	Position of D band/cm^{-1}	Position of G band/cm^{-1}	I_D/I_G	FWHM (cm^{-1})	
				D-band	G-band
GO	1352	1600	0.99	102	66
rGO	1347	1575	1.07	85	65
ZG-1.25	1343	1591	1.01	88	66
ZG-2.5	1347	1587	1.01	83	66
ZG-5	1351	1579	1.01	81	66
ZG-10	1339	1583	1.01	86	67

functional groups containing oxygen [52]. The results indicate that GO is successfully reduced to rGO after chemical reduction by hydrazine hydrate.

The Raman spectrum of all the prepared nanocomposites (Fig. 5b – e) contains the Raman bands of ZnO and rGO which confirms that the nanocomposites are successfully synthesized [53]. The I_D/I_G ratio of all the ZnO-rGO nanocomposites decreases to 1.01 compared to rGO which indicates that the sp^2 domain size of carbon atoms decreases with the incorporation of rGO in ZnO [10]. The D band shifts from 1343 to 1351 cm^{-1} and the G band decreases from 1591 to 1579 cm^{-1} as the weight percentage of rGO increases from 1.25 to 5% in ZnO. But, the same trend of shifting in the D and G band is not observed in ZG-10 nanocomposite.

A small shifting in the D and G band (Table 2) of the Raman spectrum of the nanocomposites is attributed to the electronic interactions of ZnO and rGO during the hydrothermal synthesis process [49]. In the nanocomposites, the 2D band is at $\sim$ 2700 cm^{-1} and the D + G band is at $\sim$ 2900 cm^{-1} which reconfirms the presence of rGO in the nanocomposites. The 2D and D + G bands of ZG-10 nanocomposite are not observed clearly.

To derive more quantitative information, the Raman spectra of GO (Fig. S4a in the Supplementary Information), rGO (Fig. S4b in the Supplementary Information), ZnO and ZnO-rGO nanocomposites is further deconvoluted (Fig. 6). The full-width half maximum (FWHM) is calculated using deconvoluted D and G bands (Table 2). It is found that when GO is reduced to rGO the FWHM values for both the bands decreases. The decrease in the width of the D-band of rGO is attributed to an increase in carbon-sp^2 content [54]. The lower the value of the width of G-band of rGO, higher its degree of crystallization [55]. Furthermore, with the addition of rGO (from 1.25 to 5%) in the ZnO-rGO nanocomposite, the width of the

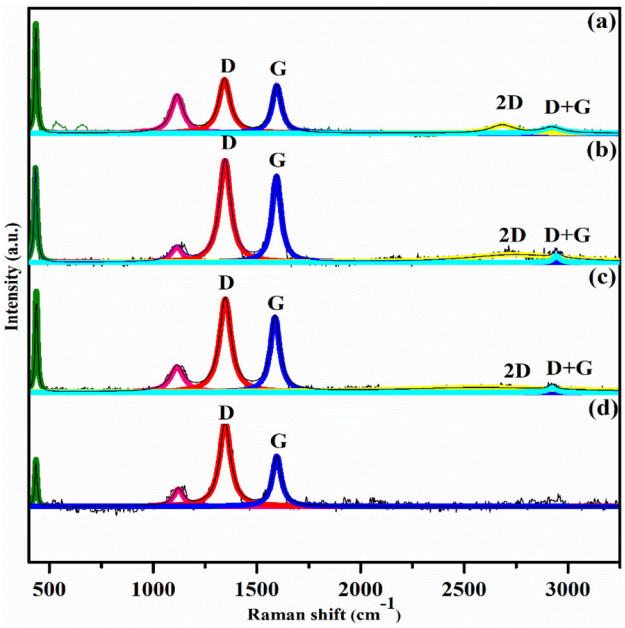


Fig. 6 The deconvoluted Raman data of ZnO and ZnO-rGO nanocomposites: a ZnO, b ZG-1.25, c ZG-2.5, d ZG-5, and e ZG-10

D-band first increases and then decreases as compared to rGO. Consequently, the carbon-sp^2 content of the ZG-5 nanocomposite is more than the ZG-1.25 nanocomposite. The width of the G-band remains almost the same in all the prepared nanocomposites. This indicates that the crystallization of rGO is maintained in all the nanocomposites. For ZG-10 nanocomposite, the width of D and G bands deviates from the trend as followed by all the other nanocomposites.

The reason for the deviation from the trend in Raman results of ZG-10 nanocomposite can be inferred as: The uncritical ratio of zinc ions to rGO does not lead to the synergetic effect between ZnO and rGO. Because of this, the ZG-10 nanocomposite cannot follow the trend as followed by the other

nanocomposites. The co-observed Raman peaks of ZnO together with rGO confirm that the ZnO-rGO nanocomposites are synthesized successfully which is in good accordance with the SEM and FTIR Spectroscopy results as well.

3.5 Electrical analysis of films

The impact of incorporation rGO on electrical properties [56, 57] of ZnO is studied using room temperature I–V plots using Keithley 2450 SMU. Figure 7 shows the illustration of I–V measurement experimental setup. Figure 8 represents the current–voltage (I–V) characteristics for Al/ZnO/ITO and Al/ZnO-rGO/ITO films with variation in weight percentage of rGO in ZnO. Figure 9a–e shows the variation in dual-logarithmic I–V curves with a varying weight percentage of rGO from 1.25 to 10% in the prepared nanocomposite films.

As can be seen from the I–V curves (Fig. 8), the forward current in the ZnO films is 51 μA at 2.5 V. When the weight percentage of rGO is increased from 1.25 to 10% in ZnO nanocomposite films (as confirmed from the EDX spectra), the forward current is found to be increasing from ≈ 51 to 105 μA (Table S1 in Supplementary Information). When rGO is incorporated from 1.25 to 5% in ZnO nanocomposite films, the leakage current under applied reverse bias voltage is found to be decreasing from 71 to 35 μA at 2.5 V. It is also found that the forward and reverse voltage in ZG-10 nanocomposite films could reach a maximum of 1.75 V after which the films breakdown. The enhancement in forward current and reduction in leakage current in the prepared nanocomposite films can be attributed to rGO nanosheets which provided a facile pathway for the direct transportation of charge carrier in the ZnO-rGO nanocomposite [58]. SEM and EDX observations, FTIR, and RAMAN

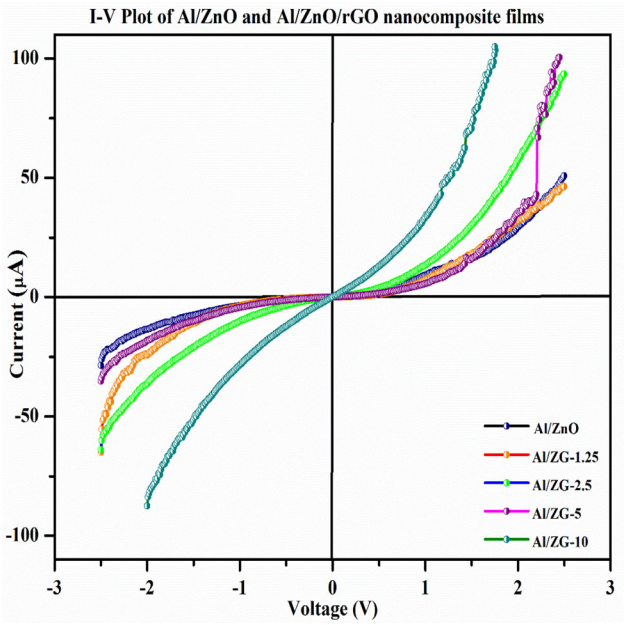


Fig. 8 Current–Voltage characteristics of ZnO and ZnO-rGO nanocomposite films

Spectroscopy analysis support these results by showcasing the coexistence of ZnO and rGO in all the prepared ZnO-rGO nanocomposites. When an external bias is applied to the films, mobility as well as the drift velocity of the charge carriers increases which thereby increased the forward current. The increase in leakage current in the ZG-10 nanocomposite films can be ascribed to the formation of trap-level defects in ZnO due to the high concentration of rGO.

The rectification ratio (RR = I_F/I_R) is calculated for all the film samples [59]. In this case, the RR is calculated as: the forward current at 2.5 V divided by the reverse current at -2.5 V for all the films except for ZG-10 nanocomposite films. The RR for ZG-10 nanocomposite films is calculated as: the forward

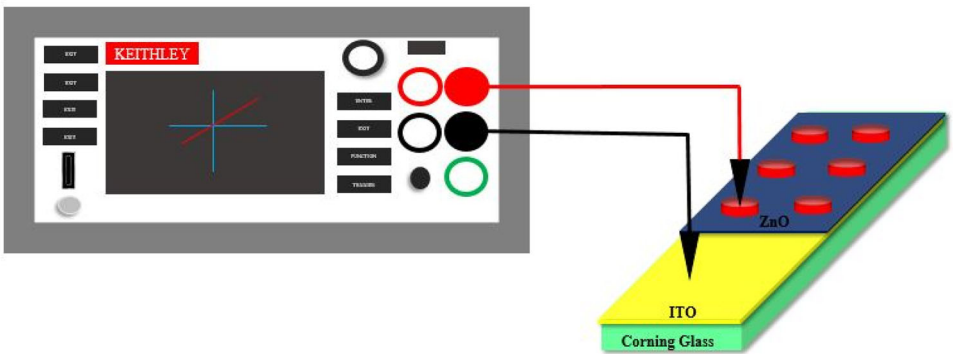


Fig. 7 Illustration of I–V measurement experimental setup

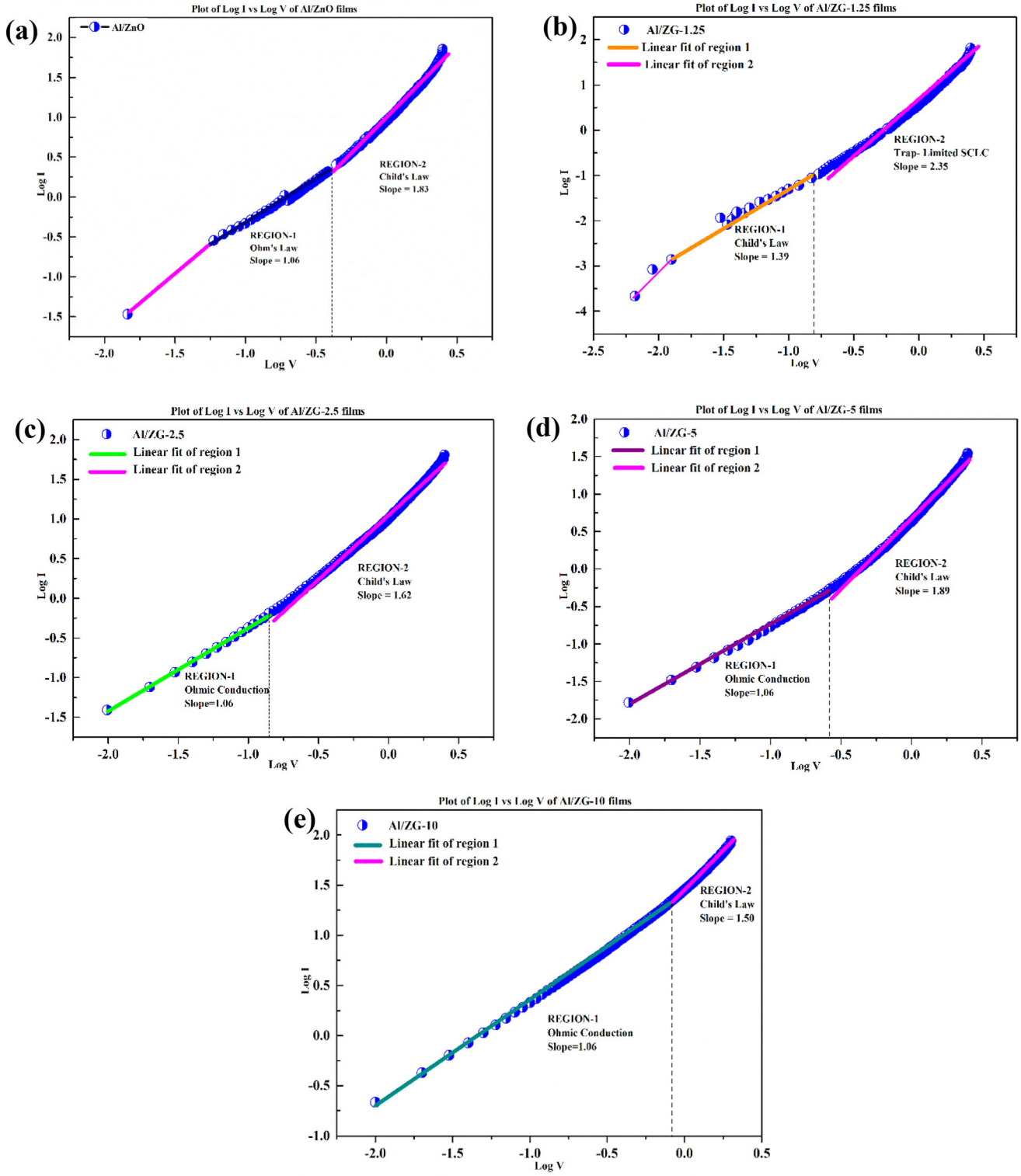


Fig. 9 Dual-logarithmic I–V curves obtained for **a** ZnO, **b** ZG-1.25, **c** ZG-2.5, **d** ZG-5, and **e** ZG-10 nanocomposite films with standard error $\leq \pm 0.002$

current at 1.75 V divided by the reverse current at − 1.75 V. The RR is found to be 0.7, 1.2, 1.5, 2.9, and 1.6 for Al/ZnO, Al/ZG-1.25, Al/ZG-2.5, Al/ZG-5, and Al/ZG-10 nanocomposite films, respectively.

The I–V curves of all the samples are almost symmetric in the forward and reverse direction, which results in low RR of Al/ZnO and Al/ZnO-rGO nanocomposite films. The reason can be attributed to the low work function contrast between the ZnO and Al electrodes, ZnO-rGO nanocomposite, and Al electrodes [60]. Table S1 in the Supplementary Information depicts the variation in electrical parameters of the fabricated ZnO and ZnO-rGO nanocomposite films with varying weight percentage of rGO from 1.25 to 10%.

To understand the effect of weight percentage variation of rGO in ZnO nanocomposite films on underlying current conduction mechanism, linear fittings of the obtained I–V curves are carried out. All the I–V curves are fitted by the space charge limited conduction (SCLC) mechanisms, which is similar to the results reported in the literature [18, 43, 61, 62]. The standard error for linear fitting of different regions in the current–voltage curves is found to be $\leq \pm 0.002$. The relationship between the current generated and applied voltages can be directly interpreted from the value of the obtained slopes from dual-logarithmic I–V curves. The obtained slopes imply that the charge transport mechanism in prepared ZnO-rGO nanocomposite films is governed by the three limiting laws of space charge limited conduction (SCLC) model, which are Ohm's law, Child's law, and trap-limited SCLC mechanism. The power-law relationship ($I \propto (V)^{m}$, where m is the slope of the plots) is used to categorize the charge transport mechanism in all the films [43]. Such a mechanism is explained based on deep-level defects. Traps capture the electronic charges, which generally lie in the forbidden gap region or extended band tails. Energetically favored states or trap states strongly influence the conduction characteristics in semiconducting materials. The reason for the creation of traps could be (1) inadvertent impurities, (2) defects introduced during the film processing [63–66]. Moreover, the addition of rGO with different weight percentage in ZnO nanocomposite films could also be the cause of interfacial defect states in the prepared films which is supported by the XRD analysis as well. The XRD analysis clearly shows an increase in dislocation density, variation in lattice parameters, and decrease in crystallite size of ZnO with the increase in weight percentage of rGO in ZnO-rGO nanocomposites. Further, EDX spectra show the presence of sodium in the as-synthesized samples which may also be one of the reasons for the creation of traps inside ZnO and ZnO-rGO nanocomposite.

Figure 9a depicts the dual-logarithmic I–V curve of ZnO films. It is observed that for low applied voltages i.e., for $V \leq 0.9$ V (Region-1), the slope of the dual-logarithmic I–V plots comes nearly one. This implies linear dependence of current on the applied voltage. Hence, it is concluded that the conduction mechanism is governed by Ohm's law. Such behavior is because the density of thermally generated free carriers inside the ZnO films is more than those of the injected charge carriers [42]. The weak injection of charge carriers from the electrodes leads to the partial filling of the trap centers. Again to maintain the neutrality of the charges inside the films, redistribution of the charge carriers occurs. This redistribution occurs in such a manner that the injected carriers are still obstructed to travel across the films which results in a slope of 1.06. At higher applied voltage ($V \geq 0.9$ V), the departure from ohmic behavior is observed thereby following the Child's Law. This is denoted by Region-2, where the electrons injected from the electrodes to the films are high. Due to which the accumulation of the charge carriers at the electrodes forms a space charge region which thereby affects the electrical distribution.

Figure 9b shows the log$I$—log$V$ curves of ZG-1.25 nanocomposite films. When rGO is introduced in the ZnO nanocomposite films, the charge carriers are increased. These charge carriers facilitate the injected charge carriers upto a certain limit to travel across the prepared films. But the effect due to traps overshadows the carriers of rGO and results in a slope of 1.39 at a low applied voltage ($V \leq 0.9$ V) indicating Child's law SCLC mechanism. A transition from Child's Law to trap-limited SCLC mechanism is observed at a higher applied voltage ($V \geq 0.9$ V). The strong injection of charge carriers at higher applied voltage occurs, due to which the carriers do not get enough transit time for reapportioning via thermally generated charge carriers or rGO charge carriers. This leads to the filling of trap centers. Hence, a slope greater than 2 is obtained which confirms the trap-limited SCLC conduction mechanism.

Figure 9c–e depicts the dual-logarithmic I–V curves of ZG-2.5, ZG-5, and ZG-10 nanocomposite films. It is observed that as the concentration of rGO is increased from 2.5 to 10% in the ZnO nanocomposite films, a clear transition of the voltage ranges for ohmic conduction keeps on increasing thereby

decreasing the voltage range for Child's Law mechanism. Apart from the prepared ZG-1.25 nanocomposite films, in which the transported charge carriers are low in comparison to the other nanocomposite films with high rGO content, all the nanocomposite films followed the same trend. The reason for such a behavior is explained based on deep-level defects as: It can be observed that for low applied voltages i.e., (Region-1), the slope of the plots comes nearly one. The films are said to follow Ohm's law. The charge carriers due to rGO incorporation increases, which not only facilitates the injected charge carriers to travel across the films but also overshadows the effect of traps. As the voltage increases, the prepared films follow Child's Law. This implies that the strong injection of charge carriers occurs due to which accumulation of the charge carriers at the electrodes form a space charge region which thereby affects the electrical distribution. Hence, the dual-logarithmic I–V characteristics of the prepared ZnO and ZnO-rGO nanocomposite films are dominated by the SCLC conduction mechanism.

4 Conclusion

Motivated by the adjoining merits of ZnO and rGO, the present study addresses the fundamental charge transport properties of ZnO and ZnO-rGO nanocomposite films and their potency for the realization of film-based nanodevices. The nanocomposites are synthesized by a facile hydrothermal method and are analysed by various characterizations. Phase formation of ZnO and dispersion of rGO in ZnO are confirmed by XRD and SEM observations. The elemental composition of ZnO and ZnO-rGO nanocomposites is confirmed by EDX analysis. The coexistence of ZnO and rGO in the nanocomposites is confirmed by FTIR and Raman spectroscopy. Films of ZnO and the ZnO-rGO nanocomposites are fabricated on ITO coated corning-glass substrate using spin-coating technique. Charge transport properties of all the as-prepared films (Al/ZnO/ITO and Al/ZnO-rGO/ITO) have been studied via current–voltage (I–V) characteristics over the voltage range of − 2.5 to 2.5 V. The detailed I–V curves shows the enhancement in the performance characteristics of the nanocomposites films by varying the rGO concentration in ZnO-rGO nanocomposite. It is found that as the rGO concentration increases in ZnO nanocomposite films, the forward current in the films also increases. Furthermore, the dual-logarithmic I–V plots have been explained based on the space charge limited conduction (SCLC) model to identify the charge transport mechanism in the films. As the weight percentage of 1.25% of rGO is introduced in ZnO nanocomposite films, the dual-logarithmic I–V characteristics give a clear transition from Child's law to trap-limited SCLC mechanism (0.9 V being the cross-over voltage). On further incorporation of rGO (from 2.5 to 10%), the conduction is favored by Ohm's law at low applied voltages and Child's law at higher applied voltages. Experimental results show that the best percentage of rGO in ZnO-rGO nanocomposite films is 5% in respect of high forward current, low leakage current, and higher rectification ratio compared to other ZnO-rGO nanocomposites. Beyond 5% of rGO in ZnO-rGO nanocomposites, the critical ratio of zinc ions to rGO is not maintained, due to which the synergetic effect did not lead to expected results. Hence, manipulation of functional properties such as charge transport by the addition of highly conducting rGO may provide additional liberty during device fabrication suitable for applications such as UV-photodetection.

Acknowledgements

The authors gratefully acknowledge the Vice-Chancellor, Prof. Yogesh Singh, Delhi Technological University, Delhi, India for his meticulous support for arranging the research facilities. The authors are grateful to Dr. Kamal Arora, Ritika Khatri, Anurag Bhandari, and Saroj Kumar Jha for their continuous support in preparing the manuscript as well as in the analysis of characterization techniques. The authors sincerely thank the Advanced Instrumentation Research Facility (AIRF), JNU, Delhi, India for providing sophisticated characterization techniques.

Declarations

Conflict of interest The authors declare that there is no conflict of interest.

Supplementary Information: The online version contains supplementary material available at https://doi.org/10.1007/s10854-021-07445-6.

References

1. R. Ramírez-Amador, J. Alvarado, G. Flores-Carrasco, L. Morales-Delagarza, S. Alcántara-Iniesta, A. Luna-Flores, Y.P. Bernal, M.Á. Méndezrojas, J.J. Gervacio-Arciniega, H.P. Martínez Hernández, J.F. Curioca-Vega, The influence of deposition time on the structural, morphological, optical and electrical properties of ZnO-rGO nanocomposite thin films grown in a single step by USP. Curr. Comput.-Aided Drug Des. **10**(2), 73 (2020). https://doi.org/10.3390/cryst10020073
2. I. Sameera, R. Bhatia, V. Prasad, R.J.A.P.L. Menon, Temperature dependent current-voltage characteristics of zinc oxide nanowire/polypyrrole nanocomposite. Appl. Phys. Lett. **105**(23), 232112 (2014). https://doi.org/10.1063/1.4903923
3. Liyana, G.R., Sofyan, N., Dhaneswara, D., Subhan, A., Yuwono, A.H.: Optoelectronic properties of ZnO nanorods thin films derived from chemical bath deposition with different growth times. In: AIP Conference Proceedings, vol. 1, p. 030008. AIP Publishing LLC. (2020) https://doi.org/10.1063/5.0015869
4. D. Nath, S. Mandal, D. Deb, J. Rakshit, P. Dey, J.J.J.O.A.P. Roy, Light tuning DC and AC electrical properties of ZnO-rGO based hybrid nanocomposite film. J. Appl. Phys. **123**(9), 095115 (2018). https://doi.org/10.1063/1.5016098
5. J. Yin, F. Gao, C. Wei, Q.J.S.R. Lu, Water amount dependence on morphologies and properties of ZnO nanostructures in double-solvent system. Sci. Rep. **4**(1), 1–7 (2014). https://doi.org/10.1038/srep03736
6. M. Morsy, I. Yahia, H. Zahran, F. Meng, M.J.J.O.E.M. Ibrahim, Portable and battery operated ammonia gas sensor based on CNTs/rGO/ZnO nanocomposite. J. Electron. Mater. **48**(11), 7328–7335 (2019). https://doi.org/10.1007/s11664-019-07550-7
7. Y. Zhao, L. Liu, T. Cui, G. Tong, W.J.A.S.S. Wu, Enhanced photocatalytic properties of ZnO/reduced graphene oxide sheets (rGO) composites with controllable morphology and composition. Appl. Surf. Sci. **412**, 58–68 (2017). https://doi.org/10.1016/j.apsusc.2017.03.207
8. Y. Liu, Y. Hu, M. Zhou, H. Qian, X.J.A.C.B.E. Hu, Microwave-assisted non-aqueous route to deposit well-dispersed ZnO nanocrystals on reduced graphene oxide sheets with improved photoactivity for the decolorization of dyes under visible light. Appl. Catal. B **125**, 425–431 (2012). https://doi.org/10.1016/j.apcatb.2012.06.016
9. S.-C. Weng, S. Brahma, C.-C. Chang, J.L.J.I.J.E.S. Huang, Synthesis of self-assembled hollow-sphere ZnO/rGO nanocomposite as anode materials for lithium-ion batteries. Int. J. Electrochem. Sci **14**, 3727–3739 (2019). https://doi.org/10.20964/2019.04.63
10. Z.K. Bolaghi, S. Masoudpanah, M.J.M.R.B. Hasheminiasari, Photocatalytic activity of ZnO/RGO composite synthesized by one-pot solution combustion method. Mater. Res. Bull. **115**, 191–195 (2019). https://doi.org/10.1016/j.materresbull.2019.03.024
11. Z. Liu, L. Yu, F. Guo, S. Liu, L. Qi, M. Shan, X.J.A.S.S. Fan, Facial development of high performance room temperature NO2 gas sensors based on ZnO nanowalls decorated rGO nanosheets. Appl. Surf. Sci. **423**, 721–727 (2017). https://doi.org/10.1016/j.apsusc.2017.06.160
12. Z. Zhan, L. Zheng, Y. Pan, G. Sun, L.J.J.O.M.C. Li, Self-powered, visible-light photodetector based on thermally reduced graphene oxide–ZnO (rGO–ZnO) hybrid nanostructure. J. Mater. Chem. **22**(6), 2589–2595 (2012). https://doi.org/10.1039/C1JM13920G
13. H. Abdullah, N.A. Atiqah, A. Omar, I. Asshaari, S. Mahalingam, Z. Razali, S. Shaari, J. Mandeep, H.J.J.O.M.S.M.I.E. Misran, Structural, morphological, electrical and electron transport studies in ZnO–rGO (wt%= 0.01, 0.05 and 0.1) based dye-sensitized solar cell. J. Mater. Sci. **26**(4), 2263–2270 (2015). https://doi.org/10.1007/s10854-015-2679-y
14. K.P. Madhuri, K. Bramhaiah, N.S.J.M.R.E. John, Nanoscale photocurrent distribution in ultra-thin films of zinc oxide nanoparticles and their hybrid with reduced graphene oxide. Mater. Res. Exp. **3**(3), 035004 (2016). https://doi.org/10.1088/2053-1591/3/3/035004
15. S. Tiagulskyi, R. Yatskiv, H. Faitova, Š Kučerová, J. Vaniš, J.J.M.S.I.S.P. Grym, Electrical properties of nanoscale pn heterojunctions formed between a single ZnO nanorod and GaN substrate. Mater. Sci. Semicond **107**, 104808 (2020). https://doi.org/10.1016/j.mssp.2019.104808
16. Y. Al-Hadeethi, R.I. Badran, A. Umar, S.H. Al-Heniti, B.M. Raffah, S.J.M.E. Al-Zhrani, Electrical properties of Ga-doped ZnO nanowires/Si heterojunction diode. Mater. Exp. **10**(6), 794–801 (2020). https://doi.org/10.1166/mex.2020.1725
17. S. Nongthombam, S. Sinha, N.A. Devi, S. Rai, R. Bhujel, W.I. Singh, B.P. Swain, Charge Transfer Mechanism of Gallium Nitrite/Reduced Graphene Oxide (GaN/rGO) Nanocomposite. In: 2020 IEEE VLSI device circuit and system (VLSI DCS) 2020, pp. 171–175. IEEE. https://doi.org/10.1109/VLSIDCS47293.2020.9179877
18. M. Soylu, M.J.J.O.A. Coskun, Compounds: controlling the properties of ZnO thin films by varying precursor concentration. J. Alloys Compds. **741**, 957–968 (2018). https://doi.org/10.1016/j.jallcom.2018.01.079
19. S.K. Mandal, K. Dutta, S. Pal, S. Mandal, A. Naskar, P.K. Pal, T. Bhattacharya, A. Singha, R. Saikh, S.J.M.C. De, Physics: Engineering of ZnO/rGO nanocomposite photocatalyst towards rapid degradation of toxic dyes. Mater. Chem.

Phys. **223**, 456–465 (2019). https://doi.org/10.1016/j.matchemphys.2018.11.002

20. M. Darvishi, F. Jamali-Paghaleh, M. Jamali-Paghaleh, J.J.S. Seyed-Yazdi, Interfaces: Facile synthesis of ZnO/rGO hybrid by microwave irradiation method with improved photoactivity. Surf. Interfaces **9**, 167–172 (2017). https://doi.org/10.1016/j.surfin.2017.09.008
21. T.N. Reddy, J. Manna, R.K.J.A.A.M. Rana, interfaces: Polyamine-mediated interfacial assembly of rGO-ZnO nanostructures: a bio-inspired approach and enhanced photocatalytic properties. ACS Appl. Mater. Interfaces **7**(35), 19684–19690 (2015). https://doi.org/10.1021/acsami.5b04820
22. D.A. Reddy, R. Ma, T.K.J.C.I. Kim, Efficient photocatalytic degradation of methylene blue by heterostructured ZnO–RGO/RuO2 nanocomposite under the simulated sunlight irradiation. Ceram. Int. **41**(5), 6999–7009 (2015). https://doi.org/10.1016/j.ceramint.2015.01.155
23. M. Ghorbani, M.R. Golobostanfard, H.J.A.S.S. Abdizadeh, Flexible freestanding sandwich type ZnO/rGO/ZnO electrode for wearable supercapacitor. Appl. Surf. Sci. **419**, 277–285 (2017). https://doi.org/10.1016/j.apsusc.2017.05.060
24. P. Sengunthar, K. Bhavsar, C. Balasubramanian, U.J.A.P.A. Joshi, Physical properties and enhanced photocatalytic activity of ZnO-rGO nanocomposites. Appl. Phys. A **126**(7), 1–9 (2020). https://doi.org/10.1007/s00339-020-03753-6
25. D.C. Marcano, D.V. Kosynkin, J.M. Berlin, A. Sinitskii, Z. Sun, A. Slesarev, L.B. Alemany, W. Lu, J.M.J.A.N. Tour, Improved synthesis of graphene oxide. ACS Nano **4**(8), 4806–4814 (2010). https://doi.org/10.1021/nn1006368
26. O.J.C. Akhavan, The effect of heat treatment on formation of graphene thin films from graphene oxide nanosheets. Carbon **48**(2), 509–519 (2010). https://doi.org/10.1016/j.carbon.2009.09.069
27. R. Bomila, A. Venkatesan, S.J.O. Srinivasan, Structural, luminescence and photocatalytic properties of pure and octylamine capped ZnO nanoparticles. Optik **158**, 565–573 (2018). https://doi.org/10.1016/j.ijleo.2017.12.141
28. H.-H. Huang, K.K.H. De Silva, G. Kumara, M.J.S.R. Yoshimura, Structural evolution of hydrothermally derived reduced graphene oxide. Sci. Rep. **8**(1), 1–9 (2018). https://doi.org/10.1038/s41598-018-25194-1
29. J. Jayachandiran, J. Yesuraj, M. Arivanandhan, A. Raja, S.A. Suthanthiraraj, R. Jayavel, D.J.J.O.I. Nedumaran, O. Polymers, Materials: Synthesis and electrochemical studies of rGO/ZnO nanocomposite for supercapacitor application. J. Inorg. Organometall. Polym. Mater. **28**(5), 2046–2055 (2018). https://doi.org/10.1007/s10904-018-0873-0
30. G. Qu, G. Fan, M. Zhou, X. Rong, T. Li, R. Zhang, J. Sun, D.J.A.O. Chen, Graphene-modified ZnO nanostructures for low-temperature NO2 sensing. ACS Omeg **4**(2), 4221–4232 (2019). https://doi.org/10.1021/acsomega.8b03624
31. K. Ravi, B.S. Mohan, G.S. Sree, I.M. Raju, K. Basavaiah, B.V.J.I.J.C.S. Rao, ZnO/RGO nanocomposite via hydrothermal route for photocatalytic degradation of dyes in presence of visible light. Int J Chem Stud **6**(6), 20–26 (2018)
32. K. Arora, S. Srivastava, P.R. Solanki, N.K.J.I.S.J. Puri, Electrochemical hydrogen gas sensing employing palladium oxide/reduced graphene oxide (PdO-rGO) nanocomposites. IEEE Sens. J. **19**(18), 8262–8271 (2019). https://doi.org/10.1109/JSEN.2019.2918360
33. B.D. Cullity, *Elements of X-ray Diffraction* (Addison-Wesley Publishing, Boston, 1956)
34. A. Goktas, A. Tumbul, Z. Aba, M.J.T.S.F. Durgun, Mg doping levels and annealing temperature induced structural, optical and electrical properties of highly c-axis oriented ZnO: Mg thin films and Al/ZnO: Mg/p-Si/Al heterojunction diode. Thin Solid Films **680**, 20–30 (2019)
35. A. Tumbul, F. Aslan, S. Demirozu, A. Goktas, A. Kilic, M. Durgun, M.Z.J.M.R.E. Zarbali, Solution processed boron doped ZnO thin films: influence of different boron complexes. Mater. Res. Exp. **6**(3), 035903 (2018). https://doi.org/10.1088/2053-1591/aaf4d8
36. H. Mohseni, H. Shokrollahi, I. Sharifi, K.J.J.O.M. Gheisari, M. Materials, Magnetic and structural studies of the Mn-doped Mg–Zn ferrite nanoparticles synthesized by the glycine nitrate process. J. Magn. Magn. Mater. **324**(22), 3741–3747 (2012)
37. D. Bobade, S. Rathod, M.L.J.P.B.C.M. Mane, Sol–gel autocombustion synthesis, structural and enhanced magnetic properties of Ni2+ substituted nanocrystalline Mg–Zn spinel ferrite. Physica B **407**(18), 3700–3704 (2012). https://doi.org/10.1016/j.physb.2012.05.017
38. P. Labhane, L. Patle, V. Huse, G. Sonawane, S.J.C.P.L. Sonawane, Synthesis of reduced graphene oxide sheets decorated by zinc oxide nanoparticles: crystallographic, optical, morphological and photocatalytic study. Chem. Phys. Lett. **661**, 13–19 (2016). https://doi.org/10.1016/j.cplett.2016.08.041
39. A. Goktas, F. Aslan, I.H.J.J.O.M.S.M.I.E. Mutlu, Annealing effect on the characteristics of La 0.67 Sr 0.33 MnO3 polycrystalline thin films produced by the sol–gel dip-coating process. J. Mater. Sci. **23**(2), 605–611 (2012)
40. A. Goktas, F. Aslan, B. Yeşilata, I.J.M.S.I.S.P. Boz, Physical properties of solution processable n-type Fe and Al co-doped ZnO nanostructured thin films: role of Al doping levels and annealing. Mater. Sci. Semicond. Process. **75**, 221–233 (2018). https://doi.org/10.1016/j.mssp.2017.11.033
41. Y.-L. Chen, Z.-A. Hu, Y.-Q. Chang, H.-W. Wang, Z.-Y. Zhang, Y.-Y. Yang, Wu, H.-Y.J.T.J.o.P.C.C., Zinc oxide/

reduced graphene oxide composites and electrochemical capacitance enhanced by homogeneous incorporation of reduced graphene oxide sheets in zinc oxide matrix. J. Phys. Chem. C **115**(5), 2563–2571 (2011). https://doi.org/10.1021/jp109597n

42. A. Göktaş, A. Tumbul, F.J.J.O.S.-G.S. Aslan, Technology: Grain size-induced structural, magnetic and magnetoresistance properties of Nd 0.67 Ca 0.33 MnO3 nanocrystalline thin films. J Sol-Gel Sci Technol **78**(2), 262–269 (2016)
43. S. Dewan, M. Tomar, R. Tandon, V.J.J.O.A.P. Gupta, Zn doping induced conductivity transformation in NiO films for realization of pn homo junction diode. J. Appl. Phys. **121**(21), 215307 (2017). https://doi.org/10.1063/1.4984580
44. X. Liu, L. Pan, Q. Zhao, T. Lv, G. Zhu, T. Chen, T. Lu, Z. Sun, C.J.C.E.J. Sun, UV-assisted photocatalytic synthesis of ZnO–reduced graphene oxide composites with enhanced photocatalytic activity in reduction of Cr (VI). Chem. Eng. J. **183**, 238–243 (2012). https://doi.org/10.1016/j.cej.2011.12.068
45. B. Li, H.J.J.O.M.C. Cao, ZnO@ graphene composite with enhanced performance for the removal of dye from water. J. Mater. Chem. **21**(10), 3346–3349 (2011). https://doi.org/10.1039/C0JM03253K
46. L. Shi, S.J.N.R.I. Gunasekaran, Preparation of pectin–ZnO nanocomposite. Nanoscale Res. Lett. **3**(12), 491–495 (2008). https://doi.org/10.1007/s11671-008-9185-6
47. N. Bano, I. Hussain, A. El-Naggar, A.J.A.P.A. Albassam, Reduced graphene oxide nanocomposites for optoelectronics applications. Appl. Phys. **125**(3), 1–7 (2019). https://doi.org/10.1007/s00339-019-2518-8
48. M.E. Uddin, R.K. Layek, N.H. Kim, D. Hui, J.H.J.C.P.B.E. Lee, Preparation and properties of reduced graphene oxide/polyacrylonitrile nanocomposites using polyvinyl phenol. Compos. B **80**, 238–245 (2015). https://doi.org/10.1016/j.compositesb.2015.06.009
49. S. Kumar, V. Pandit, K. Bhattacharyya, V.J.M.C. Krishnan, Physics: Sunlight driven photocatalytic reduction of 4-nitrophenol on Pt decorated ZnO-RGO nanoheterostructures. Mater. Chem. Phys. A **214**, 364–376 (2018). https://doi.org/10.1016/j.matchemphys.2018.04.113
50. C. Rodwihok, D. Wongratanaphisan, Y.L. Thi Ngo, M. Khandelwal, S.H. Hur, J.S.J.N. Chung, Effect of GO additive in ZnO/rGO nanocomposites with enhanced photosensitivity and photocatalytic activity. Nanomaterials **9**(10), 1441 (2019). https://doi.org/10.3390/nano9101441
51. R. Beura, P.J.J.O.P. Thangadurai, C.O. Solids, Structural, optical and photocatalytic properties of graphene-ZnO nanocomposites for varied compositions. J. Phys. Chem. Solids **102**, 168–177 (2017). https://doi.org/10.1016/j.jpcs.2016.11.024
52. H. Fan, X. Zhao, J. Yang, X. Shan, L. Yang, Y. Zhang, X. Li, M.J.C.C. Gao, ZnO–graphene composite for photocatalytic degradation of methylene blue dye. Catal. Commun. **29**, 29–34 (2012). https://doi.org/10.1016/j.catcom.2012.09.013
53. Y.T. Li, J.M. Xu, Z.J. Tang, T.T. Xu, X.J.J.J.O.A. Li, Compounds: Nearly white light photoluminescence from ZnO/rGO nanocomposite prepared by a one-step hydrothermal method. J. Alloys Compds. **715**, 122–128 (2017). https://doi.org/10.1016/j.jallcom.2017.04.286
54. X. Díez-Betriu, S. Álvarez-García, C. Botas, P. Álvarez, J. Sánchez-Marcos, C. Prieto, R. Menéndez, A.J.J.O.M.C.C. De Andrés, Raman spectroscopy for the study of reduction mechanisms and optimization of conductivity in graphene oxide thin films. J. Mater. Chem. C **1**(41), 6905–6912 (2013)
55. H. Liu, M. Xiang, X.J.J.O.M.S. Shao, Graphene/ZnO nanocomposite with seamless interface renders photoluminescence quenching and photocatalytic activity enhancement. J. Mater. Sci. **53**(19), 13924–13935 (2018). https://doi.org/10.1007/s10853-018-2605-9
56. S. Demirezen, A. Kaya, Ö. Vural, ŞJ.M.S.I.S.P. Altındal, The effect of Mo-doped PVC+ TCNQ interfacial layer on the electrical properties of Au/PVC+ TCNQ/p-Si structures at room temperature. Mater. Sci. Semicond. Process **33**, 140–148 (2015). https://doi.org/10.1016/j.mssp.2015.01.050
57. S.A. Yerişkin, M. Balbaşı, I.J.J.O.M.S.M.I.E. Orak, The effects of (graphene doped-PVA) interlayer on the determinative electrical parameters of the Au/n-Si (MS) structures at room temperature. J. Mater. Sci. Mater. Electron. **28**(18), 14040–14048 (2017). https://doi.org/10.1007/s10854-017-7255-1
58. S. Safa, R. Sarraf-Mamoori, R. Azimirad, The effects of reduced graphene oxide (rGO) on ZnO film UV-detector. In: Advanced Materials Research 2014, pp. 577–582. Trans Tech Publ. https://doi.org/10.4028/www.scientific.net/AMR.829.577
59. A. Eroğlu, S. Demirezen, Y. Azizian-Kalandaragh, S.S.J.J.O.M.S.M.I.E. Altındal, A comparative study on the electrical properties and conduction mechanisms of Au/n-Si Schottky diodes with/without an organic interlayer. J. Mater. Sci **31**(17), 14466–14477 (2020). https://doi.org/10.1007/s10854-020-04006-1
60. M. Mohiuddin, B. Kumar, S.J.B.C.I.E. Haque, Biopolymer composites in photovoltaics and photodetectors. Biopolym. Compos. Electron. **45**, 459–486 (2017). https://doi.org/10.1016/B978-0-12-809261-3.00017-6
61. B. Murali, S.J.D.T. Krupanidhi, Transport properties of CuIn 1–x Al x Se 2/AZnO heterostructure for low cost thin film photovoltaics. Dalton Trans. **43**(5), 1974–1983 (2014). https://doi.org/10.1039/C3DT52515E

62. X. Yan, H. Hao, Y. Chen, S. Shi, E. Zhang, J. Lou, Liu, B.J.N.R.l., Self-rectifying performance in the sandwiched structure of Ag/In-Ga-Zn-O/Pt bipolar resistive switching memory. Nanoscale Res. Lett. **9**(1), 1–6 (2014). https://doi.org/10.1186/1556-276X-9-548
63. E. Tanrıkulu, S. Demirezen, Ş Altındal, İJ.J.O.M.S.M.I.E. Uslu, Analysis of electrical characteristics and conduction mechanisms in the Al/(% 7 Zn-doped PVA)/p-Si (MPS) structure at room temperature. J Mater Sci **28**(12), 8844–8856 (2017). https://doi.org/10.1007/s10854-017-6613-3
64. S. Demirezen, S.A.J.P.B. Yerişkin, A detailed comparative study on electrical and photovoltaic characteristics of Al/p-Si photodiodes with coumarin-doped PVA interfacial layer: the effect of doping concentration. Polym. Bull. **77**(1), 49–71 (2020). https://doi.org/10.1007/s00289-019-02704-3
65. S. Demirezen, H. Çetinkaya, M. Kara, F. Yakuphanoğlu, ŞJ.S. Altındal, A.A. Physical, Synthesis, electrical and photosensing characteristics of the Al/(PCBM/NiO: ZnO)/p-Si nanocomposite structures. Sens. Actuators A **317**, 112449 (2021). https://doi.org/10.1016/j.sna.2020.112449
66. Chiu, F.-C., Pan, T.-M., Kundu, T.K., Shih, C.-H.: Thin film applications in advanced electron devices. In. Hindawi, (2014). https://doi.org/10.1155/2014/927358

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)
Pune, India. Oct 29-30, 2021

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON) | 978-1-6654-2503-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/SMARTGENCON51891.2021.9645922

Link Prediction in Social Networks using Histogram Based Gradient Boosting Regression Tree

Himanshu Tiwari
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
himanshutiw2794@gmail.com

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

***Abstract*— Social Networks play an important role in today's world and applications range from a connected world to uncovering important relationships in bio-logical systems. The enormous rise in the application of social networks has resulted in the importance of detecting the relationships between the nodes even before they are formed. Several approaches have been proposed in the past specifically for the task of link prediction using various indices. There has been a tremendous effort in merging various indices using machine learning concepts and analogy of Law of Gravitation using similarity measures as a proxy for distance and popularity measures as a proxy of mass. The merging of several indices can increase the overall effectiveness of link prediction and very few algorithms have been suggested in past. In this work, we propose the use of a "Histogram based Gradient Boosting Regression Tree" for the task of link prediction after unifying three popularity and four similarity measures.**

Keywords—link prediction, node embeddings, social networks, complex network, Histogram-based Gradient Boosting Regression Tree

I. INTRODUCTION

The recent developments in the field of epidemiology have resulted in an in-creased focus on social interactions and the study of social networks. Social me-dia is getting popular among all age groups and societies resulting in enormous amounts of creation of data and growth of the underlying social networks. The study of social networks has always been a subject of fascination. Social Net-works are often represented by people and the relationships between them. As a general rule, social networks are composed of nodes and edges, where people are the nodes and relationships are their edges. Social networks, of course, are quite dynamic. In a normal social network, new interactions and nodes are formed at a breakneck speed. The relationships evolve over time and may result in deepening or even the creation of new relationships over time. Social networks are useful for a variety of purposes, including information transmission, event prediction, and the prediction of new relationships. Because of the huge scale of the structures, the majority of tasks linked to social networks are algorithmically hard. Although most of the time the social networks deal with the relationships formed among social groups, they have found a huge number of applications in other fields as well like molecular biology. Insights to the information social networks can also provide meaningful and vast knowledge about community influence maximization [1, 2], detection [3], and information diffusion [3].

Link prediction is one such problem in the field of social networks. It deals with detecting relationships among nodes before even they are formed or they appear in the perceived social network. Link prediction is the process of predicting associations before they form. This effort is critical because we can accurately forecast numerous aspects of networks, such as the recommendations of friends on social media, the contributions of scholarly writers, and even the sub-terranean links between terrorists. Using link prediction techniques, we may even forecast or discover relationships between the many participants in a given act. These techniques can also be utilized for customized product recommendations on e-commerce websites and personalized online marketing.

For the link prediction problem, numerous algorithms have been developed over time, including maximum likelihood-based link prediction algorithms, similarity-based link prediction algorithms, and probability-based link prediction algorithms [5]. The concept behind similarity approaches is that the greater the similarity (having multiple common properties) of nodes, the better the likelihood of future link development [4]. The networks are built on some assumptions and principles, and maximum likelihood approaches are based on those assumptions and principles [6, 7]. Detailed rules are established based on these assumptions and principles, and the possibility of a currently non-existing relationship forming in the future is assessed. Most networks are arranged in hierarchies on various levels, according to these assumptions. These approaches are mostly dependent on network structure properties such as hierarchies or communities. Probability-based approaches look for a function that abstracts the underlying network and fits the data to generate an observed network. The future links can be anticipated using conditional probability and the parameters that the probabilistic function is based on once this model has been established.

In the recent past, certain algorithms influenced by physical laws have also been described [8, 9, 10]. These methods largely rely on physical formulas and the analogy of various measures in the context of social networks. The Law of Gravitation formula is the most widely used formula. The gravitational force be-tween two masses is proportional to the product of their masses and inversely proportional to the square of the distance between them, according to this formula. The researchers adopted the analogy of portraying the bodies under consideration as network nodes, with gravitational force serving as the link prediction task's metric.

Single similarity indexes are unable to provide such exact findings, therefore, few researchers have also introduced the idea of a merger, however, there are only a few approaches

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:00:05 UTC from IEEE Xplore. Restrictions apply.

that use this technique. The merging of several indices can increase the overall effectiveness of link prediction, according to existing merger-related research. This is because it can get through the problem of a single index only being suitable for networks with identical structural properties. The existing research has also established that the merging of indexes is not the primary reason for improved outcomes; rather, it is also the selection of all those indices that affects link prediction results [11, 12].

We suggested a method called Histogram-based Gradient Boosting Regression Tree for Link Prediction (HGBLP) based on the fusion of three popularity and four similarity measures based on the aforesaid study. The method appears to yield superior results in theory, as various indices represent different network features.

The remainder of the paper is organized in the following manner. Section II discusses related work in the field of link prediction and perhaps even the problem statement, while Section III discusses our suggested solution, " Histogram based Gradient Boosting Regression Tree for Link Prediction (HGBLP)" The results and evaluation criteria are discussed in Section IV, and the conclusion and future scope of this project are discussed in Section V.

II. Related Work

The link prediction challenge entails predicting the emergence of new linkages over time. In more formal terms, imagine we have a snapshot of a social network at current time *t*, and we want to reliably anticipate the edges that are not present at time t but are there at time *t+h*.

Similarity-based techniques assign a score to a pair of nodes that reflect a non-existent link. The number of common neighbors between them can be used to calculate similarity scores. This strategy has been investigated in a number of research works, with a positive association between link construction and greater common neighbor values. For each pair of nodes x and y, link prediction algorithms based on similarity provide them a similarity score. This score is based on how similar or close x and y are [13, 14]. For link prediction, maximum likelihood algorithms have just been presented in the recent past. By studying the structures and determining the highest likelihood, these algorithms are able to obtain specific rules and parameters. The criteria and parameters are used to anticipate new linkages. Even with a few thousand nodes, these methods are typically exceedingly sluggish. Various probability-based models have been pro-posed in the past. The Relational Bayesian Network, Probabilistic Entity-Relationship Models (PERMs), and Stochastic Relational Models are some of the examples of Probability Relational Models (PRMs) [15, 16]. Several writers have also proposed several nature-inspired link prediction techniques [8, 9, 10]. These methods mostly rely on physics rules and attempt to incorporate multiple net-work measurements into the formula. A gravity law-based link production model was proposed by Bastami, Mahabad, and Taghizadeh. For the reduction of the prediction space, they used community and network information. To replace the mass and distances in the gravitation law, Wahid-Ul-Ashraf, Budka, and Musial employed a variety of combinations for similarity and centrality of nodes. In Newton's Gravitation Law, they employed and evaluated a combination of three node centrality measurements as mass and thirteen dissimilarity values as distance. In the past, Node Embeddings have also been used for the task of link pre-diction [17, 18]. Node Embeddings are the low dimensional vector representation of nodes of the social network which can also be used in various machine learning algorithms. The advantage of using such techniques is that they abstract the underlying network.

Merger-based techniques tend to take up multiple similarity indices and devise a formula taking into account the effect of all these measures. These provide a performance boost since a single similarity measure works best on a limited type of network topologies. In the past fusion-based methods have been proposed including the Adaboost based [11] and Ensemble-based Link Prediction algorithms [12]. We propose a method called Histogram-based Gradient Boosting Regression Tree for Link Prediction (HGBLP) based on the fusion of three popularity and four similarity measures based on the aforesaid study.

III. Proposed Method

A. Popularity measures used

We propose using the following three popularity measures:

1) Degree Centrality: Degree centrality is the same as a node's degree in the context of graph theory. In other words, degree centrality is the number of nodes directly connected to the node in the network. This is a straightforward yet very helpful popularity index because in a way it precisely gives us local information about a network node's popularity.

2) Closeness Centrality: This index measures a node's average distance i.e. how well connected on an average that particular node is to all other nodes in the network. The inverse of the sum of shortest paths from that node to all other nodes is used to calculate this value.

3) Page Rank: This is a centrality metric that is based on the network's random walk. Consider an imaginary traveler who is moving from one node to another picking his next node randomly. This centrality measures the likelihood that this traveler will arrive at a particular node in consideration.

B. Similarity Measures used

We also propose the use of the following five similarity measures in the algorithm:

1) Common Neighbours: This is a local similarity statistic based on the number of mutual friends shared by the node pair in question. Consider two nodes having set of neighbors, the intersection of these sets represents the common neighbors.

2) Adamic Adar: This is an extended local similarity metric that is based on the number of mutual neighbors also taking into account the degrees of the nodes.

3) Cosine Similarity: Cosine similarity of two vectors represent the cosine of the angles between the them. In network theory, if the nodes are represented as vectors and the cosine of the angle between them is taken, it forms an interesting similarity measure. It can be represented as

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:00:05 UTC from IEEE Xplore. Restrictions apply.

Closeness centrality divided by the square root of the product of the degrees of the corresponding two nodes.

4) Matrix Forest Index: MFI [19] is defined as the ratio of the number of spanning forests such that v_i and v_j are contained in the same tree with v_i as the root to all spanning rooted forests of the main graph. The spanning subgraph may not have all of the graph's edges, but it does contain all of the graph's vertices.

C. Methodology

The steps involved in the proposed method are given below:

1) Reducing the network: An original network *G(V,E)* containing a set of vertices *V* and a set of edges E is used to generate a new network *G'(V,E')*. G' contains all of the vertices from the original graph G, and E' is a proper subset of *E*. This assures that some of the links in the newly established network *G*' are absent. We try to forecast the network G using the network G' for link prediction.

2) Computing the popularity measures: The newly established network is used to calculate the popularity measures for each node in the network. The popularity measures mentioned above are calculated and they will form the input for the regressor.

3) Computing the similarity indices: The similarity indices mentioned above are calculated and stored in a vector. The similarity measures are calculated for all the node pairs: $(V \times V) - E'$. The node pairs can be classified into two categories, one of them representing the edges present in the original graph but not in the new graph representing G' and the other category representing the edges that were never present in G. The task is essentially predicting the edges of first category accurately.

4) Training and Prediction: The data created in the step 2 and 3 is down-sampled and split in 7:3 ratio for training and testing. We apply *Histogram Based Gradient Boosting Regression Tree* regressor to train our model and then test it to predict the relationships.

IV. RESULTS AND DISCUSSION

A. Datasets used

The below datasets were used for the analysis of our proposed solution. These datasets are freely available, and the following is a description of them:

1) Jazz Musicians: Jazz performers represent the nodes, while an edge symbolises two musicians who have collaborated [20].

2) Les Misérables: The characters in Les Misérables are represented by this data [20].

3) Football: The network of American Football games played amongst Division IA institutions during the regular season in Fall 2000 is shown by this data [21].

4) Dolphin: Bottlenose dolphins form a social network by interacting with each other. The associations between dolphins are represented by the links in this data [22].

5) US-Air: It has 332 airports and 2126 routes, with each node representing an airport and each edge indicating whether or not a direct flight exists between two airports.

The data was preprocessed to provide a standard format with integer nodes. We didn't need to care for weights because the edges are unweighted. Table I summarises the common characteristics of the datasets mentioned previously.

TABLE I. DATASETS USED FOR EVALUATION OF THE PROPOSED SOLUTION

Dataset	Number of vertices	Number of edges	Average Degree
Jazz Musicians	198	2742	27.6970
Les Misérables	77	254	6.5974
Football	115	613	10.6609
Dolphin	62	158	5.0968
US-Air	332	2126	12

B. Evaluation Metrices

We used the area under the receiver operating characteristic (ROC) curve to assess the performance of our Histogram based Gradient Boosting Regression Tree for Link Prediction (HGBLP) technique. AUC is a well-known metric that represents the likelihood of assigning a lower score to a randomly chosen non-existent connection than to a randomly given missing link. Let's say we do N separate comparisons of our algorithm's derived scores, and out of those N comparisons, missing connections get a higher score N' times. In addition, if the score is less than or equal N'' times, the AUC value is defined by (1).

$$AUC = \frac{N' + 0.5N''}{N} \quad (1)$$

Table II represents the calculated AUC values of our proposed Histogram-based Gradient Boosting Regression Tree for Link Prediction (HGBLP) method. The algorithm has shown a good performance boost in general on the datasets evaluated in the analysis.

TABLE II. CALCULATED AUC VALUES FOR THE PROPOSED HISTOGRAM-BASED GRADIENT BOOSTING REGRESSION TREE FOR LINK PREDICTION (HGBLP) METHOD.

Dataset	Calculated AUC Value
Jazz Musicians	0.888650453
Les Misérables	0.919090909
Football	0.787076023
Dolphin	0.826190476
US-Air	0.898901617

Table III represents the calculated AUC values for the method proposed by the authors in literature [9] using the nature-inspired approach. In the literature, they suggested the use of the combination of Closeness centrality and Matrix Forest Index (MFI) for best results. We used the method suggested by the authors using the same parameters for the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:00:05 UTC from IEEE Xplore. Restrictions apply.

calculation of the AUC values on the aforementioned datasets.

It is evident from the results that our proposed solution outperformed the nature-inspired approach on each dataset used. The results are compared in Fig. 1.

TABLE III. CALCULATED AUC VALUES USING THE NATURE INSPIRED APPROACH - CCM.

Dataset	Calculated AUC Value
Jazz Musicians	0.522205232
Les Misérables	0.704840614
Football	0.566232998
Dolphin	0.630699088
US-Air	0.5862818

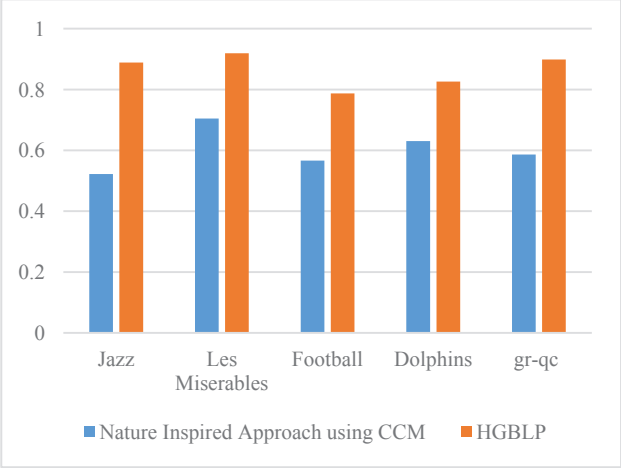


Fig. 1. Comparison of the AUC values on different datasets using the Nature Inspired Approach CCM and our proposed HGBLP solution.

We also compared and evaluated our solution against the Matrix Forest Index similarity measure [19]. The calculated AUC values for the Matrix Forest Index link prediction task are shown in Table IV. When compared with our solution, it is evident that our solution performed better on the used datasets on each instance (see Fig. 2).

TABLE IV. CALCULATED AUC VALUES USING MFI (MATRIX FOREST INDEX) FOR LINK PREDICTION.

Dataset	Calculated AUC Value
Jazz Musicians	0.506762996
Les Misérables	0.725754765
Football	0.570047309
Dolphin	0.569040382
US-Air	0.511013286

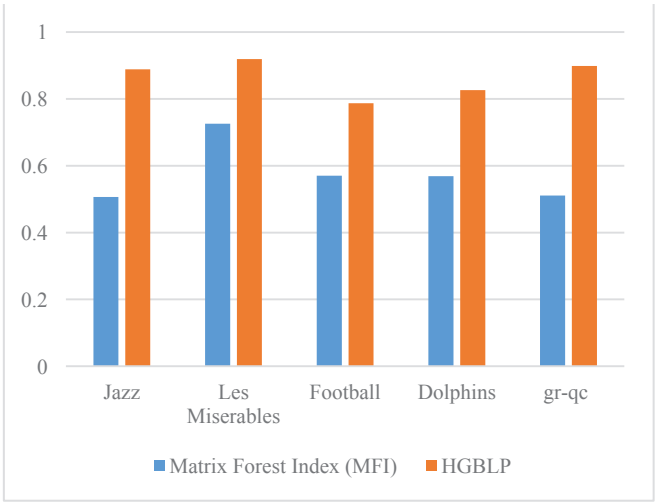


Fig. 2. Comparison of the AUC values on different datasets using the Matrix Forest Index (MFI) and our proposed HGBLP solution.

Finally, we evaluated our solution against the previously proposed fusion-based Ensemble Based Link Prediction (EMLP) algorithm defined in the literature [12]. The calculated AUC values for the link prediction task using the EMLP solution are shown in Table V.

TABLE V. CALCULATED AUC VALUES USING EMLP (ENSEMBLE MODEL BASED LINK PREDICTION).

Dataset	Calculated AUC Value
Jazz Musicians	0.864144785
Les Misérables	0.869090909
Football	0.747076023
Dolphin	0.726190476
US-Air	0.826321688

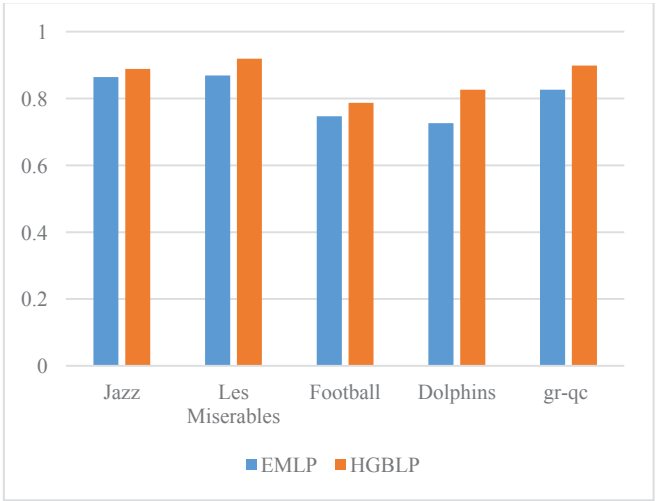


Fig. 3. Comparison of the AUC values on different datasets using the EMLP and our proposed HGBLP solution.

It is evident from Fig. 3 that our proposed solution HGBLP outperforms the previous fusion-based solution on the datasets mentioned in Table 1.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:00:05 UTC from IEEE Xplore. Restrictions apply.

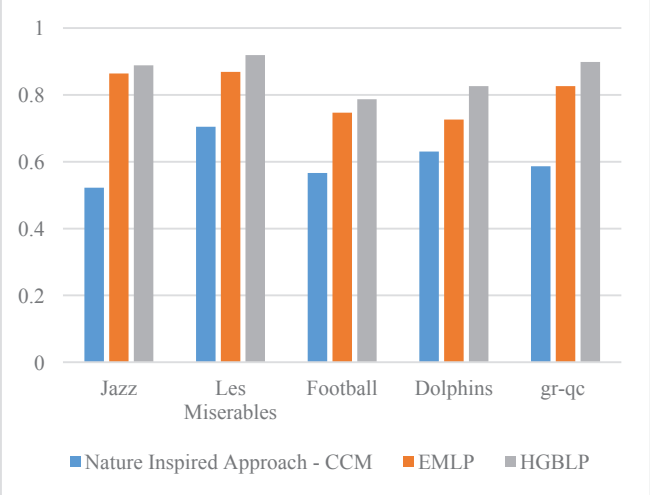


Fig. 4. Comparison of the AUC values on different datasets using all the fusion-based algorithms

The comparison of all the fusion-based solutions in Fig. 4 also suggests that the proposed HGBLP solution outperforms the previous fusion-based solutions. The nature-inspired approach fused one similarity and one popularity measure in the form of gravitational law formula.

V. Conclusion

The fusion-based solutions like the like EMLP and Nature-inspired solutions generally outperform the traditional single similarity-based solutions because they tend to average out the effect of a particular topology and structural proper-ty working in favor or against a particular measure. Finding the structural proper-ties in an evolving social network is a challenging task and fusion-based solutions might be a good choice in such cases. Our proposed solution, HGBLP works better than the previously proposed methods in the literature.

References

[1] Kumar, S., Singhla, L., Jindal, K. et al. IM-ELPR: Influence maximization in social networks using label propagation based community structure. Appl Intell (2021). https://doi.org/10.1007/s10489-021-02266-w

[2] Kumar, S., Lohia, D., Pratap, D. et al. MDER: modified degree with exclusion ratio algorithm for influence maximisation in social networks. Computing (2021). https://doi.org/10.1007/s00607-021-00960-8

[3] Kumar S., Kumar M. (2019) Predicting Customer Churn Using Artificial Neural Network. In: Macintyre J., Iliadis L., Maglogiannis I., Jayne C. (eds) Engineering Applications of Neural Networks. EANN 2019. Communications in Computer and Information Science, vol 1000. Springer, Cham. https://doi.org/10.1007/978-3-030-20257-6_25

[4] D. Lin, An information-theoretic definition of similarity, in Proceedings of the 15th International Conference on Machine Learning, Morgan Kaufman Publishers, San Francisco, 1998.

[5] Lü, Linyuan, and Tao Zhou. "Link Prediction in Complex Networks: A Survey." Physica A: Statistical Mechanics and its Applications 390.6 (2011): 1150–1170. Crossref. Web.

[6] M. Sales-Pardo, R. Guimer`a, L. A. N. Amaral, Extracting the hierarchical organization of complex systems, Proc. Natl. Acad. Sci. U.S.A. 104 (2007) 15224.

[7] M. Girvan, M. E. J. Newman, Community structure in social and biological networks, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 7821.

[8] Bastami, E., Mahabad, and Taghizadeh. 2019. ": A gravitation-based link prediction approach in social networks. Swarm and evolutionary computation." 176-186.

[9] Wahid-Ul-Ashraf, A., Budka, M., and Musial, K. 2019. "How to predict social relationships—Physics-inspired approach to link prediction." Physica A: Statistical Mechanics and its Applications, no. 523, 1110-1129.

[10] Kumar S., Chaudhary U., Kedia R., Singhal T. (2021) Link Prediction in Complex Network: Nature Inspired Gravitation Force Approach. In: Gupta D., Khanna A., Bhattacharyya S., Hassanien A.E., Anand S., Jaiswal A. (eds) International Conference on Innovative Computing and Communications. Advances in Intelligent Systems and Computing, vol 1166. Springer, Singapore. https://doi.org/10.1007/978-981-15-5148-2_28

[11] Z.F. WU , Q. Liang , Q. Liu , Z.G. Qin , Modified link prediction algorithm based on AdaBoost, J. Commun. 35 (2014) 116-123

[12] Li, K., Tu, L., & Chai, L. (2020). Ensemble-model-based link prediction of complex networks. Computer Networks, 166, 106978

[13] P. Holme, M. Huss, Role-similarity based functional prediction in networked systems: application to the yeast proteome, J. R. Soc. Interface 2 (2005) 327.

[14] G. Kossinets, Effects of missing data in social networks, Social Networks 28 (2006) 24.

[15] D. Heckerman, C. Meek, D. Koller, Probabilistic Entity-Relationship Models, PRMs, and Plate Models, In Proceedings of the 21st International Conference on Machine Learning, Banff, Canada, 2004, p. 55.

[16] K. Yu, W. Chu, S. Yu, V. Tresp, Z. Xu, Stochastic Relational Models for Discriminative Link Prediction, In Proceedings of Neural Information Precessing Systems, MIT Press, Cambridge MA, 2006, p. 1553.

[17] Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2016.

[18] Mingdong Ou, Peng Cui, Jian Pei, Ziwei Zhang, and Wenwu Zhu. 2016. Asymmetric Transitivity Preserving Graph Embedding. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). Association for Computing Machinery, New York, NY, USA, 1105–1114. DOI:https://doi.org/10.1145/2939672.2939751

[19] P. Chebotarev, E. Shamis, The matrix-forest theorem and measuring relations in small social groups (2006) arXiv preprint math/0602070.

[20] "SNAP Datasets." n.d. SNAP Datasets. https://snap.stanford.edu/data.

[21] "American College Football Dataset." n.d. http://www.casos.cs.cmu.edu/computational_tools/datasets/external/football/index2.html.

[22] "Dolphins Dataset." n.d. http://networkrepository.com/soc-dolphins.php.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 07:00:05 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.

Metasurface Superstrate Beam Steering Antenna with AMC for 5G/WiMAX/WLAN Applications

Akhilesh Verma
Delhi Technological University
Ravi Kumar Arya (raviarya@gmail.com)
National Institute of Technology Delhi
Srinivasa Nallanthighal Raghava
Delhi Technological University

Research Article

Keywords: Monopole, Beam steering, Metasurface superstrate, 5G.

Posted Date: December 8th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-695261/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

A beam-steering antenna based on non-uniform metasurface superstrate and AMC, operating at 3.5 GHz, is presented. The antenna can steer the beam along θ = -18° and 18° with the superstrate and along θ = 0° in the absence of the superstrate with almost zero scan loss. Antenna structure consists of a top layer of non-uniform metasurface superstrate made of a 20 × 20 grid of electrically-small square-shaped metallic pixels while the bottom part consists of AMC with a grid of 5 × 5 pixels. The radiating element, CPW-fed monopole antenna, is placed between AMC and superstrate. The fabricated prototype shows desired beam steering in directions of θ = -18°, 0°, and 18° while maintaining uniform realized gain of 5.5 dB and matches well with simulations.

I. Introduction

Fifth generation (5G) is the recent wireless technology that is the next step towards the evolution of cellular communication systems [1], [2]. In 5G, the beam steering capability of the antenna plays the main role for efficient coverage to all users [3]. Beam steering is the solution to today's network problems such as call drops, multipath fading in crowded areas [4], [5]. Optical beam scanning for short range communication is presented in [6], which achieved 10° beam scanning. Beam scanning is achieved with the help of waffle and waffle iron waveguides. Many beam steering methods have been discussed and reported in [7], which shows one of the methods of beam steering is Birefringent Prisms. However, this method is limited as it requires prism thickness of more than 8 cm which may create the problem of beam walkof [8], [9] when the beam needs to be steered by 20°. The precise beam steering is obtained using 1D optical phased array on silicon-on-insulator which achieves 2.3° at wavelength of 1550 nm and 14.1° steering. The design is, however, limited due to its high cost and complication in designing with low beam scanning [10]. The reconfigurable partial reflecting surface (PRS) structure was deployed to achieve beam steering from − 15° to 15° as reported in [11]. Besides this, the microstrip leaky wave antenna has achieved frequency scanning capability with two tapered slots in the leaky section and reduction of side lobe level is obtained by Yagi elements [12]. Also, 0° to 40° beam scanning is reported in [13] with the help of an electronically controlled microstrip antenna, and changing the capacitance value. The tilted beam at ± 15° is observed in [14], by combining parasitic elements with inserted capacitances. The parasitic phased array antenna with varactors and tuning capacitance which provides continuous scanning of ± 15° is reported in [15]. A reconfigurable microstrip Yagi antenna is presented in [16] and beam scanning from − 40° to + 40° in H-plane is achieved by two hexagonal slotted tunable parasitic elements with the gain of 3.36 dBi. The beam steering of -30°, 0°, + 30° is reported in [17], which used a grid of 5 × 5 electrically-small rectangular-shaped metallic pixels on the upper layer of the parasitic layer which are connected by switches. However, 360° beam steering is presented in [18] by combining electronically steerable parasitic array radiator (ESPAR) antenna and the water antenna, but it makes the design complex and cannot be easily designed. In [19], phase varying metamaterial is used for compact steerable directive antenna which achieves a ± 20° steered beam. Furthermore, different beam steering methods have been reported in [20].

In next sections, we provide a step-by-step approach to design the proposed antenna. Sec. II starts with the design of a radiating element, CPW-fed monopole antenna. Once, radiating element is designed, the design of AMC and superstrate is discussed. Next, in Sec. III, various parameters of composite antenna with AMC and superstrate are studied so that change in antenna behavior due to change in these parameters can be quantified. This section also discusses desired radiation behaviors of the proposed antenna. Finally, Sec. IV concludes the work with concluding remarks.

Ii. Geometry Of The Proposed Antenna

(a) Design of CPW fed monopole antenna

The radiating element, CPW fed monopole antenna, geometry is shown in Fig. 1. Due to simple design and single side design on substrate, this radiating element is selected for this research. For design, the length of the monopole antenna is denoted as *l*, and width is denoted by *W*. The monopole antenna is fed through 50 Ω SMA connector. The antenna is fabricated on FR4 ($\epsilon_r = 4.4$, $\tan\delta$=0.02) substrate with a thickness of 1.6 mm. The other dimensions of the antenna are given as:
$L_p = 35,\ W_p = 22,\ l = 16.7,\ m = 9.1,\ g = 0.3,\ l_1 = 10(alldimensionsareinmm)$.

(b) Design of AMC for CPW fed monopole antenna

The center frequency of AMC operation is chosen to be 3.5 GHz. A unit cell of periodicity 10 mm × 10 mm is used for AMC design. Square-shaped metallic patches of side '*a*' are printed at the center of the unit cell on FR4 substrate as shown in Fig. 2 (a). The substrate has thickness, h_1 = 3.2 mm with permittivity of 4.4, and loss tangent of 0.02. For analysis of AMC, the periodic boundary condition (PBC) [19, 20, 21, 22, 23] is used. It facilitates quick results instead of simulating the large structure of AMC. By using PBC, shown in Fig. 2(b), the reflection phase is calculated by simulation with the help of Floquet port excitation.

In order to ascertain that the AMC is indeed working on desired 3.5 GHz, unit cell simulation with variable side of the square pixel, *a*, is run and an attempt is made so that the reflection phase is 0° at desired 3.5 GHz. From Fig. 2(c), we can see that when the side of the pixel, *a*, is 9.6 mm, the reflection phase is 0°. Henceforth, we will use AMC made of unit cells with *a* = 9.6 mm to use with the monopole antenna discussed earlier.

(c) Design of composite CPW fed monopole antenna with AMC and superstrate

In the previous section, we designed AMC. Next, we place AMC underneath the antenna substrate as shown in Fig. 3 so that the AMC enhances the radiation in forward direction, i.e., along vertically up direction. It can be seen from Fig. 3 that the grid of 5×5AMC unit cells is placed under the monopole antenna. It is well known that the superstrate can be used to increase the gain of the antenna [24, 25, 26, 27]. Mostly, the superstrate used for enhancing the design is either in the form of uniform pixels or pixels whose size changes radially out from the center of the superstrate. In such applications, the primary role of superstrate is to work in Fabry-Perot configuration and enhance the gain of the antenna element. In this research, we have a dual purpose of the superstrate. On one hand we want to modify the gain of the antenna and on the other, we want to steer the beam emanating from the antenna in the desired direction. This type of beam steering is very important for upcoming 5G systems. Figure 4 (a) shows the geometry of the proposed superstrate which comprises of metallic patches of different size from $a_1$to a_{20}. The dimensions of the proposed superstrate are given in Table 1.

In order to steer the beam in direction away from the broadside, there should be phase variation for waves emanating from different locations of the superstrate. This phase-varying behavior is provided by a nonuniform metasurface printed on the dielectric substrate acting as a superstrate. For this purpose, the size of the unit cells of the superstrate increases monotonically from one edge to another of the superstrate as shown in Fig. 4 (a). This way, the superstrate will help to steer the beam in an off-broadside direction. If this superstrate sends the beam in + θ, rotating the superstrate by 180°, taking the z-axis for rotation, will help to send the beam in the -θ direction. In order to send the beam in broadside direction, the superstrate is removed. To ease the rotation, the interchange of the superstrate may be accomplished by a mechanical solution. In this way, this design can be said to be reconfigurable. However, the speed of beam steering will depend upon how quickly the mechanical motor can help to rotate or remove the superstrate.

Figure 4 (b) shows the layered structure of the proposed CPW fed monopole antenna with AMC and superstrate with different views. The fabricated prototype of the proposed antenna is shown in Fig. 5.

Table 1
Patch dimensions (in mm) used in metasurface superstrate.

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
0.125	0.25	0.375	0.5	0.625	0.75	0.875	1	1.125	1.25
a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}	a_{19}	a_{20}
1.375	1.5	1.625	1.75	1.875	2	2.125	2.25	2.375	2.5

Iii. Results And Discussion

All the simulations in this work are conducted using Ansys HFSS, full wave simulator. The simulated S-parameter comparison of the monopole antenna, monopole antenna with AMC, monopole antenna with AMC and superstrate is shown in Fig. 6.

The comparison of the realized gain for monopole antenna with AMC and superstrate is shown in Fig. 7. It can be observed that there is 3.84 dBi increment in gain after applying AMC on the monopole antenna. Furthermore, when the superstrate is applied, the gain of the antenna is 5.5 dBi at 3.5 GHz.

A. Parametric analysis of the antenna

The proposed antenna is fed by CPW as shown in Fig. 1. The parametric studies of the dimensions of the monopole antenna, gap between antenna and superstrate has been carried out to achieve good impedance matching and desired frequency band.

i) Length of monopole antenna, l

The length of the monopole antenna, *l*, is varied from 20 mm to 30 mm with the step size of 5 mm. As shown in Fig. 8 (a) the reflection coefficient S_{11} is lower than − 10 dB for desired frequency bands. It can also be observed that the S_{11} resonance moves to lower frequency as the length of the monopole increases.

ii) Gap between the antenna and superstrate

The gap between antenna and superstrate is varied from 4.2 mm to 10.2 mm with the step size of 3 mm. It can be seen from Fig. 8. (b) that there is hardly any effect of varying the gap on the antenna performance. In this way, the antenna will give almost the same results even if the gap changes due to slight mishandling of the antenna.

B. Near field analysis

In order to get insights of working of the superstrate, near field behavior is very helpful [25]. Far field pattern being the Fourier transform of the near-field pattern, can give better understanding of how the superstrate is changing the near field that will eventually change the far field characteristics. To accomplish it, a vertical plane in the center of the antenna is placed on which the dominant component of the near field is studied. In our case, E_x is the dominant electric field as the monopole is aligned along the x-axis. In Fig. 9 and Fig. 10, E_x magnitude and phase behavior in absence and presence of the superstrate are shown. It is evident from Fig. 9 (a) and Fig. 10 (a) that the placement of the superstrate confines the fields and hence shows slightly higher field magnitude in case of the superstrate. It is also evident from Fig. 9 (b) and Fig. 10 (b) that due to placement of the superstrate, the direction of equi-phase front changes. The phase front is changed towards left, signifying that the outward radiation is expected towards left side when the superstrate is placed. From these figures, it is easy to understand that the absence or presence of the superstrate decides the beam steering direction.

After getting the desired radiation properties, the prototype of the antenna is fabricated as shown in Fig. 5. The proposed antenna is tested using R & S ZNB40 vector network analyzer and anechoic chamber for its properties.

C. S-parameter measurement

The simulated and measured S-parameters of the antenna are shown in Fig. 11. For acceptable performance, the proposed antenna needs to provide good reflection coefficients, |S11|, in the desired frequency band. From Fig. 11, it is evident that the fabricated antenna has |S11| < -10 dB over the frequency band of 3.1–3.7 GHz. The simulated and measured S-parameters are in good agreement. The fabricated antenna has wide impedance bandwidth for 5G/WiMAX/WLAN applications.

D. Radiation pattern of the proposed antenna

Figure 12 shows the radiation pattern of the proposed antenna for φ = 0° cut. The non-uniform superstrate was designed to steer the beam in θ=±18°. The superstrate needs to be mechanically rotated or removed depending upon the desired steering beam direction. While, currently, we manually did the operation of mechanically placing and removing the superstrate but the superstrate can be moved by mechanical motor to automate this operation. This type of mechanical movement of a superstrate is simple and can be easily accomplished by mechanical motors.

It can be seen from Fig. 12 that the beam scans in ± 18° and 0° in elevation directions in presence of superstrate and in absence respectively. It is also interesting to note that the antenna does not show sidelobes in this plane. The realized gains at these steered angles are 5.5 dB signifying that the antenna shows almost zero scan loss for ± 18° scanning angles.

The proposed antenna radiation patterns (co- and cross-polarization) are measured in an anechoic chamber at 3.5 GHz as shown in Figs. 13 and 14. In these figures, co-polarization is shown by red lines and cross-polarization by black lines. It is evident from Figs. 13 and 14, that the antenna has very low cross-polarization components. The simulated and measured results are also in good agreement.

E. Realized gain

Figure 15 shows the simulated and measured gain of the antenna with and without superstrate. It is observed that the gain of the antenna in both cases is 5.5 dBi at 3.5 GHz. The simulated and measured values are also in good agreement.

Comparison table

In Table 2, all existing designs are compared with the current work. It is evident from Table 2 that the proposed antenna gives a good beam steering with compact design.

Table 2
Comparison of present work with the previously published antennas

Ref	Antenna type	Size (λ^3_0)	Frequency band (GHz)	Realized gain (dBi)	Beam scanning (Degrees)
[11]	Phased array and reconfigurable PRS structure	$3.1 \times 3.1 \times 0.55$	5.5 to 5.7	12	$\pm 15^o$
[14]	Parasite patches and lumped capacitances	$1.56 \times 0.63 \times 0.013$	2.45	5–8	$\pm 15^o$
[15]	Parasitic phased array antennas	$0.30 \times 0.78 \times 0.005$	1	7.4	$\pm 15^o$
[16]	Tunable parasitic	$0.49 \times 0.49 \times 0.0013$	2.43–2.47	3.36	$\pm 40^o$
This Work	Metasurface monopole Antenna	$0.58 \times 0.58 \times 0.16$	3.1–3.7	5.53	$\pm 18^o$

Iv. Conclusion

A beam-steering antenna based on non-uniform metasurface superstrate and AMC with CPW-fed monopole antenna as radiating element operating at 3.5 GHz (3.1–3.7 GHz) was presented. The antenna is able to steer beams in directions of -18°, 0° and − 18° in elevation while maintaining uniform gain of 5.5 dBi. The antenna is compact, has a good front to back ratio and low sideline levels. The antenna due to these qualities is an ideal candidate for 5G, WiMAX and WLAN applications.

Declarations

Funding: The authors did not receive support from any organization for the submitted work

Conflicts of interest/Competing interests: The Authors declares that there is no conflict of interest'.

Availability of data and material (data transparency): Data sharing not applicable -no new data generated, or the article describes entirely theoretical research data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Code availability (software application or custom code): No codes are used in this manuscript.

Authors' contributions: All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Akhilesh Verma, Ravi Kumar Arya and Srinivasa Nallanthighal Raghava. The first draft of the manuscript was written by Akhilesh Verma and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

References

1. Yu, B., Yang, K. and Yang, G., 2017. A novel 28 GHz beam steering array for 5G mobile device with metallic casing application. *IEEE Transactions on Antennas and Propagation*, *66*(1), pp.462-466.
2. Zhang, S., Syrytsin, I. and Pedersen, G.F., 2018. Compact beam-steerable antenna array with two passive parasitic elements for 5G mobile terminals at 28 GHz. *IEEE Transactions on Antennas and Propagation*, *66*(10), pp.5193-5203.
3. Shah, A.S.A., Abd Rahman, N.H., Ali, M.T., Fauzi, N.F. and Ab Aziz, A.A., 2016, December. Beam scanning of phased array antenna using phase modification method for satellite application. In *2016 IEEE Asia-Pacific Conference on Applied Electromagnetics (APACE)* (pp. 291-295). IEEE.
4. Serhsouh, I., Himdi, M., Lebbar, H. and Vettikalladi, H., 2020. Reconfigurable SIW antenna for fixed frequency beam scanning and 5G applications. *IEEE Access*, *8*, pp.60084-60089.
5. Wu, L., Farrall, A.J. and Young, P.R., 2015. Substrate integrated waveguide switched beam antenna. *IEEE Transactions on Antennas and Propagation*, *63*(5), pp.2301-2305.
6. Arai, H., 2019, November. High gain optical beam scanning antenna and its measurement. In *2019 IEEE Asia-Pacific Conference on Applied Electromagnetics (APACE)* (pp. 1-3). IEEE.
7. McManamon, P.F., Bos, P.J., Escuti, M.J., Heikenfeld, J., Serati, S., Xie, H. and Watson, E.A., 2009. A review of phased array steering for narrow-band electrooptical systems. *Proceedings of the IEEE*, *97*(6), pp.1078-1096.
8. U. Schmidt and W. Hust, BOptical deflection system including an alternating sequence of birefringent prisms and polarizers,[U.S. Patent 3 572 895, Dec. 1986.
9. M. A. Karim, D. Cook, and P. F. McManamon, BDigital beam steering system using modules of ferro electric liquid crystal and Wollaston prism,[in Tech. Dig. 1988 Annu. Meeting Opt. Soc. Amer., Santa Clara, CA, Oct. 1988, p. 118.
10. Van Acoleyen, K., Bogaerts, W., Jágerská, J., Le Thomas, N., Houdré, R. and Baets, R., 2009. Off-chip beam steering with a one-dimensional optical phased array on silicon-on-insulator. *Optics letters*, *34*(9), pp.1477-1479.
11. Ji, L.Y., Guo, Y.J., Qin, P.Y., Gong, S.X. and Mittra, R., 2015. A reconfigurable partially reflective surface (PRS) antenna for beam steering. *IEEE Transactions on Antennas and Propagation*, *63*(6), pp.2387-2395.
12. Tanoli, S.A.K., Khan, M.I., Fraz, Q., Yang, X. and Shah, S.A., 2018. A compact beam-scanning leaky-wave antenna with improved performance. *IEEE Antennas and Wireless Propagation Letters*, *17*(5), pp.825-828.

13. Symeonidou, A. and Siakavara, K., 2013. A novel microstrip antenna array with metamaterial-based electronic beam steering at 2.4 GHz. *Progress In Electromagnetics Research*, *38*, pp.27-42.
14. Dias, M.H.C., Freitas, V. and Vuong, T.P., 2015. Low cost electronic beam tilting of microstrip antennas by the use of tunable parasites. *Microwave and Optical Technology Letters*, *57*(11), pp.2710-2713.
15. Luther, J.J., Ebadi, S. and Gong, X., 2012. A microstrip patch electronically steerable parasitic array radiator (ESPAR) antenna with reactance-tuned coupling and maintained resonance. *IEEE Transactions on Antennas and Propagation*, *60*(4), pp.1803-1813.
16. Khairnar, V.V., Kadam, B.V., Ramesha, C.K. and Gudino, L.J., 2018. A reconfigurable parasitic antenna with continuous beam scanning capability in H-plane. *AEU-International Journal of Electronics and Communications*, *88*, pp.78-86.
17. Li, Z., H. Mopidevi, O. Ğ. U. Z. Kaynar, and B. A. Cetiner. "Beam-steering antenna based on parasitic layer." *Electronics Letters* 48, no. 2 (2012): 59-60.
18. Xing, L., Zhu, J., Xu, Q., Yan, D. and Zhao, Y., 2019. A circular beam-steering antenna with parasitic water reflectors. *IEEE Antennas and Wireless Propagation Letters*, *18*(10), pp.2140-2144.
19. Ourir, A., S. N. Burokur, and A. De Lustrac. "Phase-varying metamaterial for compact steerable directive antennas." *Electronics letters* 43, no. 9 (2007): 493-494.
20. Uchendu, I. and Kelly, J.R., 2016. Survey of beam steering techniques available for millimeter wave applications. *Progress In Electromagnetics Research B*, *68*, pp.35-54.
21. Ravi Kumar Arya, Shiyu Zhang, Shaileshachandra Pandey, Ashwani Kumar, Y. Vardaxoglou, William Whittow and Raj Mittra, “Meta-Atoms and Artificially Engineered Materials for Antenna Applications,” in Developments in Antenna Analysis and Design: Volume 1 edited by Raj Mittra, pp. 351-405, IET, UK, 2018 DOI: 10.1049/SBEW543F_ch10
22. Akhilesh Verma, Ravi Kumar Arya, Rajarshi Bhattacharya, Srinivasa Nallanthighal Raghava, “Compact PIFA Antenna with High Gain and Low SAR using AMC for WLAN/C-band/5G Applications ,” *IETE Journal of Research,* 2021 doi:10.1080/03772063.2021.1945958
23. R. K. Arya, C. Pelletti and R. Mittra, “Numerically Efficient Technique for Metamaterial Modeling,” *Progress In Electromagnetics Research*, Vol. 140, pp. 263-276, 2013 DOI: 10.2528/PIER13051313
24. Anil Kumar Yerrola, Ravi Kumar Arya, Maifuz Ali, and Lakhindar Murmu, “Non-Omnidirectional Monopole using Holographic Surfaces with Reduced SLL.” URSI GASS 2021, Rome, Italy, 28 August – 4 September, 2021.
25. Li, Y., Mittra, R., Zeng, B., Lu, G., Li, Z., Liu, J., ... & Chang, D. C. (2012). Directivity enhancement of fabry-perot antenna by using a stepped-dielectric slab superstrate. Microwave and Optical Technology Letters, 54(3), 711-715.
26. Mittra, Raj, Yanfei Li, and Kyungho Yoo. "A comparative study of directivity enhancement of microstrip patch antennas using three different superstrates." Microwave and Optical technology letters 52.2 (2010): 327-331.
27. Ji, L.-Y., Guo, Y. J., Qin, P.-Y., Gong, S.-X., & Mittra, R. (2015). *A Reconfigurable Partially Reflective Surface (PRS) Antenna for Beam Steering. IEEE Transactions on Antennas and Propagation, 63(6),*

2387–2395.

28. Schneider, John B. "Understanding the finite-difference time-domain method." *School of electrical engineering and computer science Washington State University* 28 (2010).

Figures

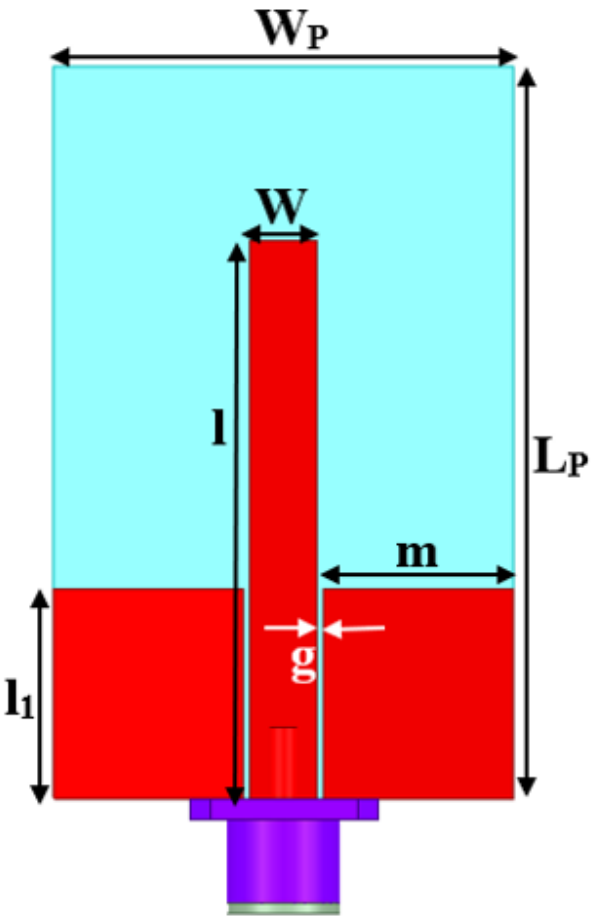


Figure 1

Proposed CPW fed monopole antenna.

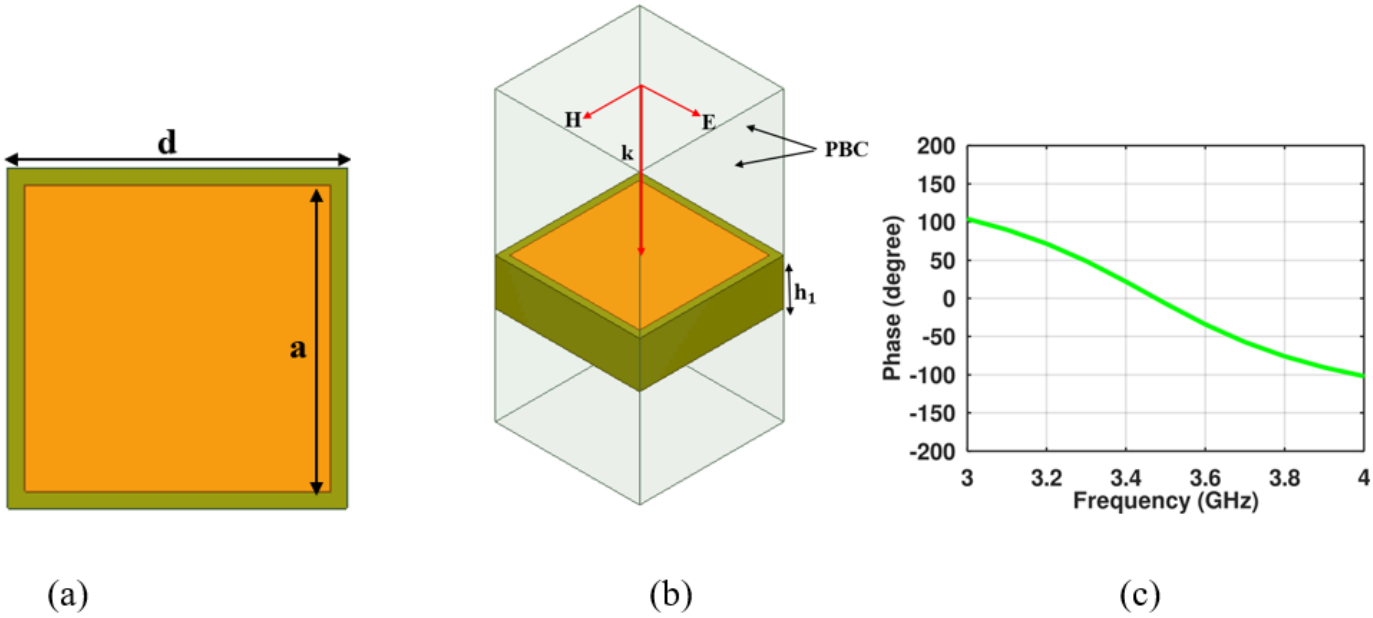


Figure 2

AMC design for proposed CPW fed monopole antenna, (a) top view of AMC, (b) isometric view of AMC, (c) reflection phase behavior of the AMC.

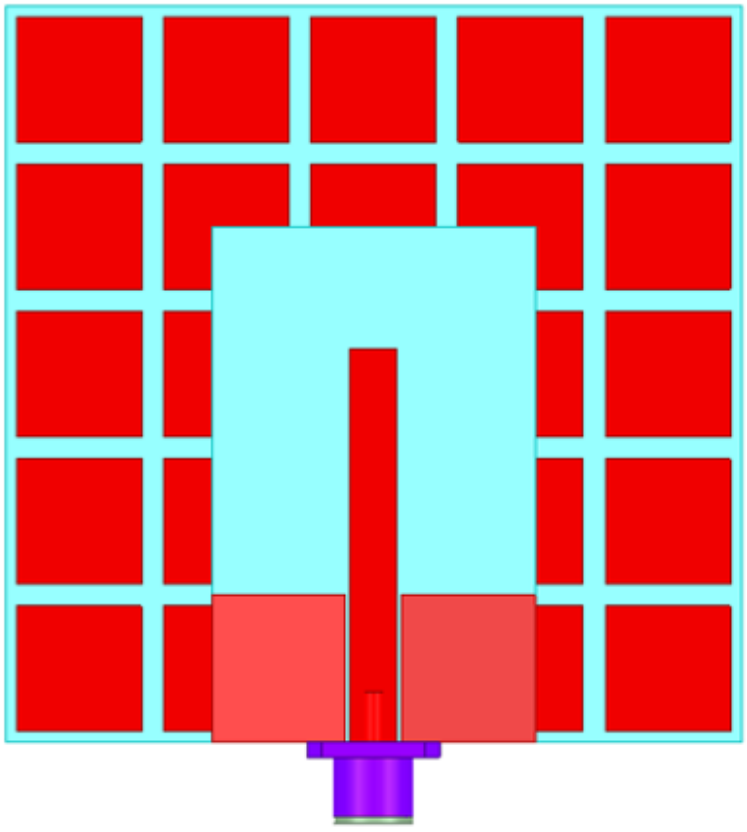


Figure 3

CPW fed monopole antenna with AMC.

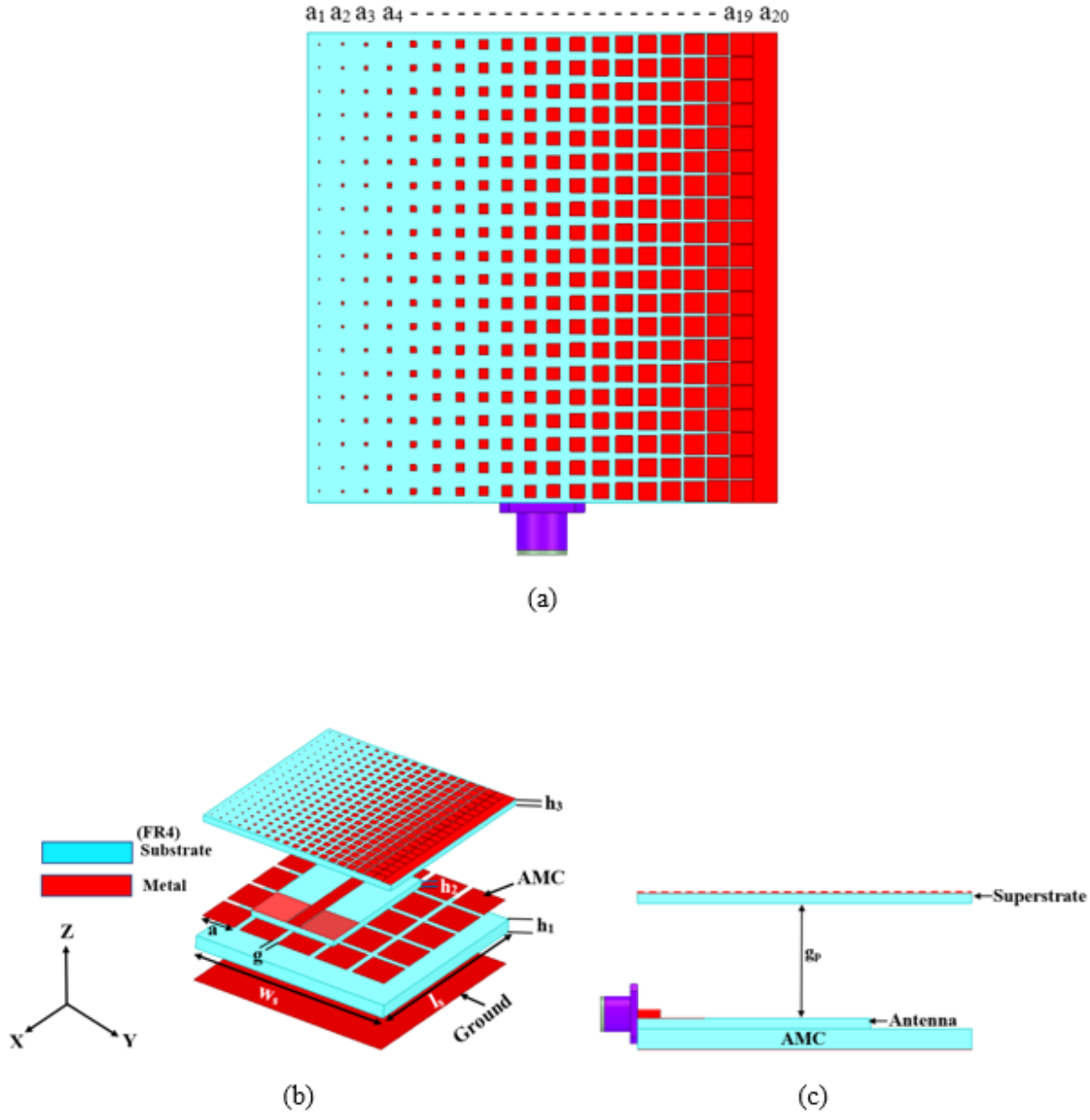


Figure 4

CPW fed monopole antenna with AMC and superstrate. (a) layered structure, (b) side view. (l_s=50,W_s=50,h_1=3.2,h_2=1.6,h_3=1.6,g=0.17,g_p=7.2 (all dimensions in mm).

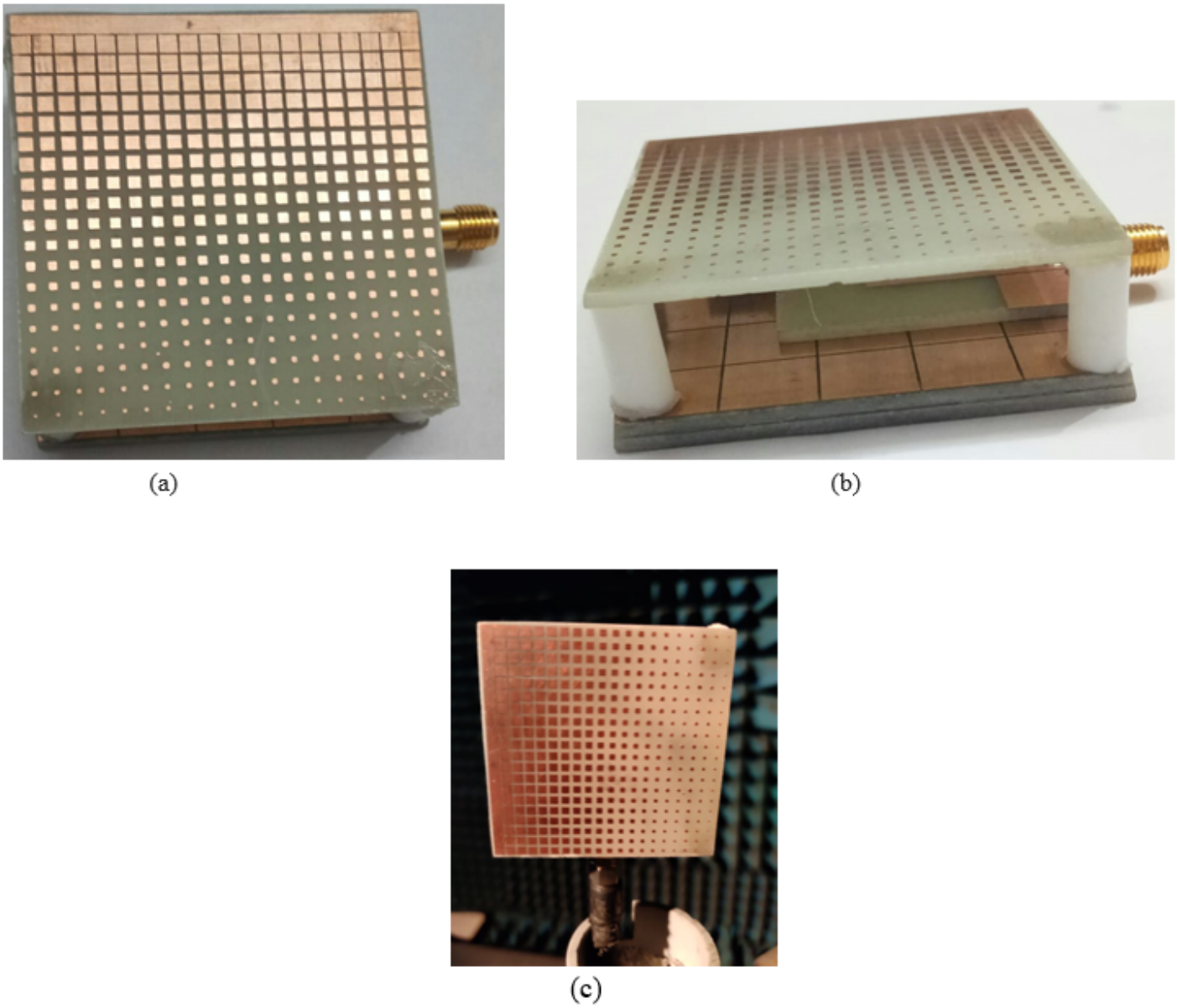


Figure 5

Prototype of the proposed CPW fed antenna with AMC and superstrate, (a) top view, (b) side view, (c) antenna in an anechoic chamber for measurement.

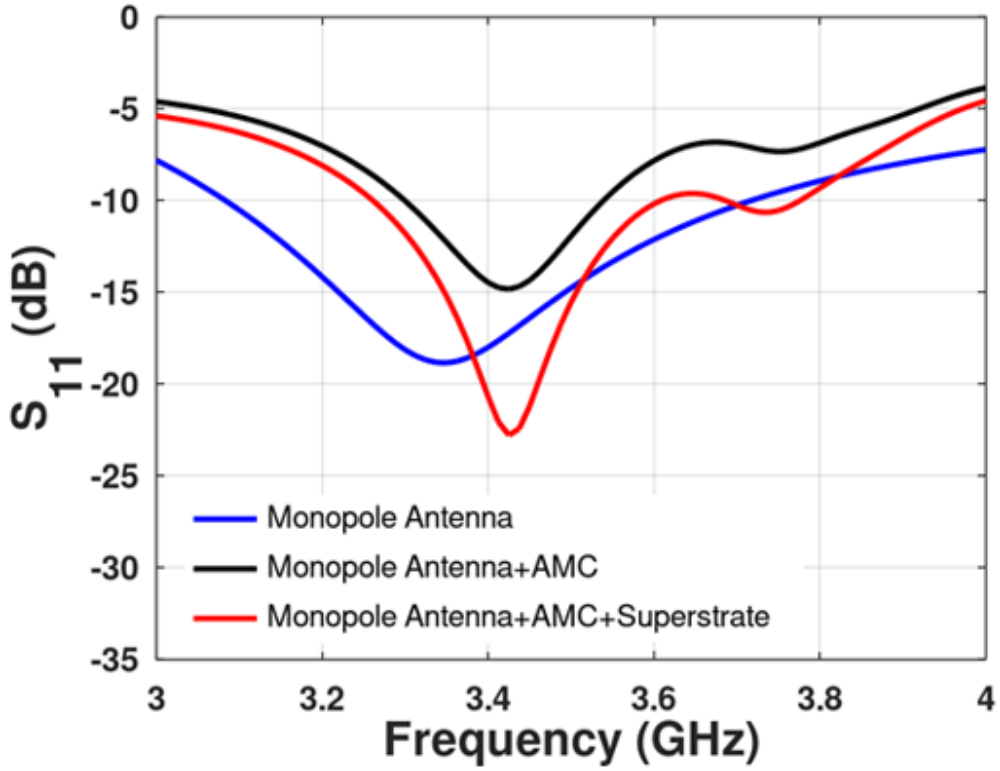


Figure 6

S-parameter comparison of the monopole antenna with AMC and superstrate.

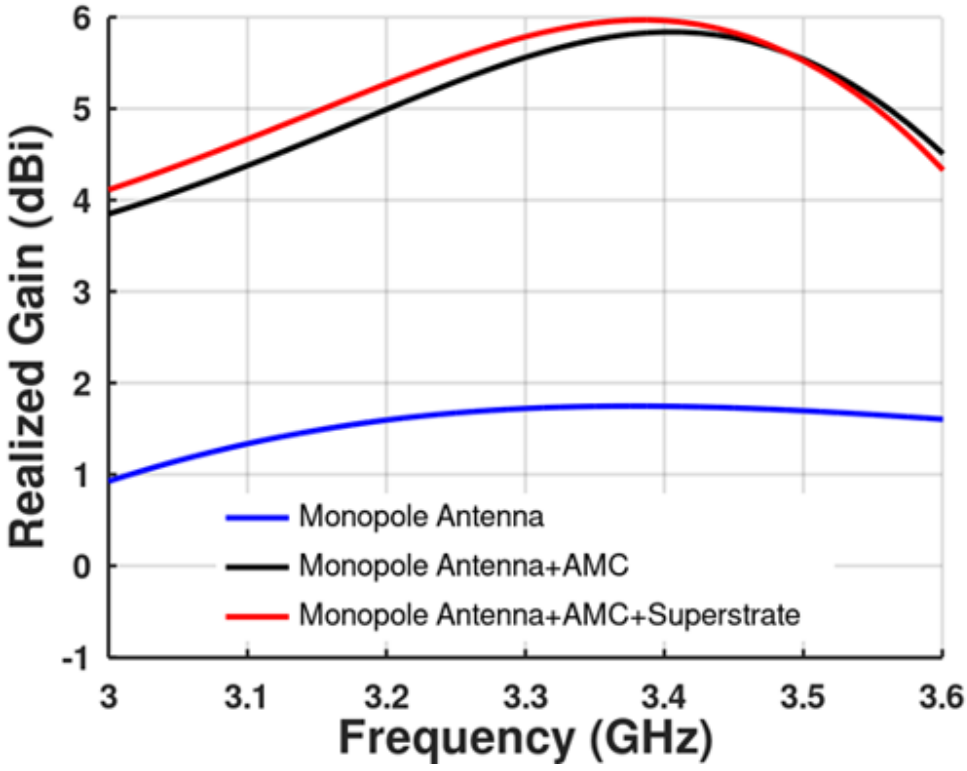


Figure 7

Realized gain comparison of the monopole antenna with AMC and superstrate.

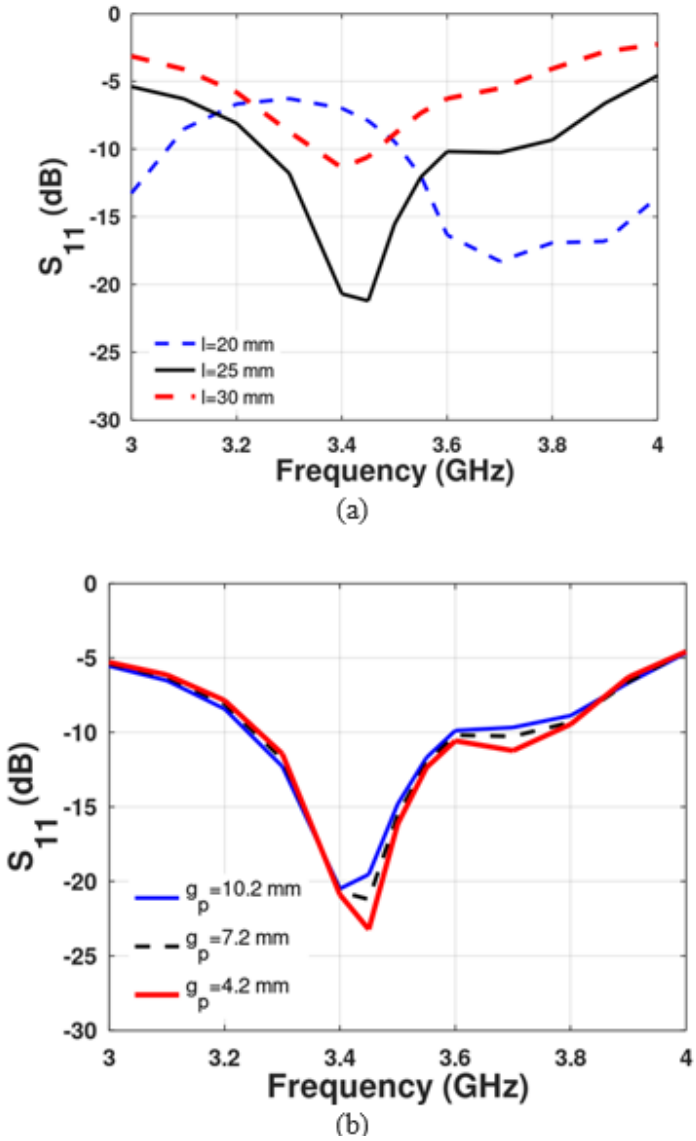


Figure 8

Parametric analysis of the antenna using: (a) monopole antenna length ‘l’; (b) gap between antenna and superstrate ‘g_p’.

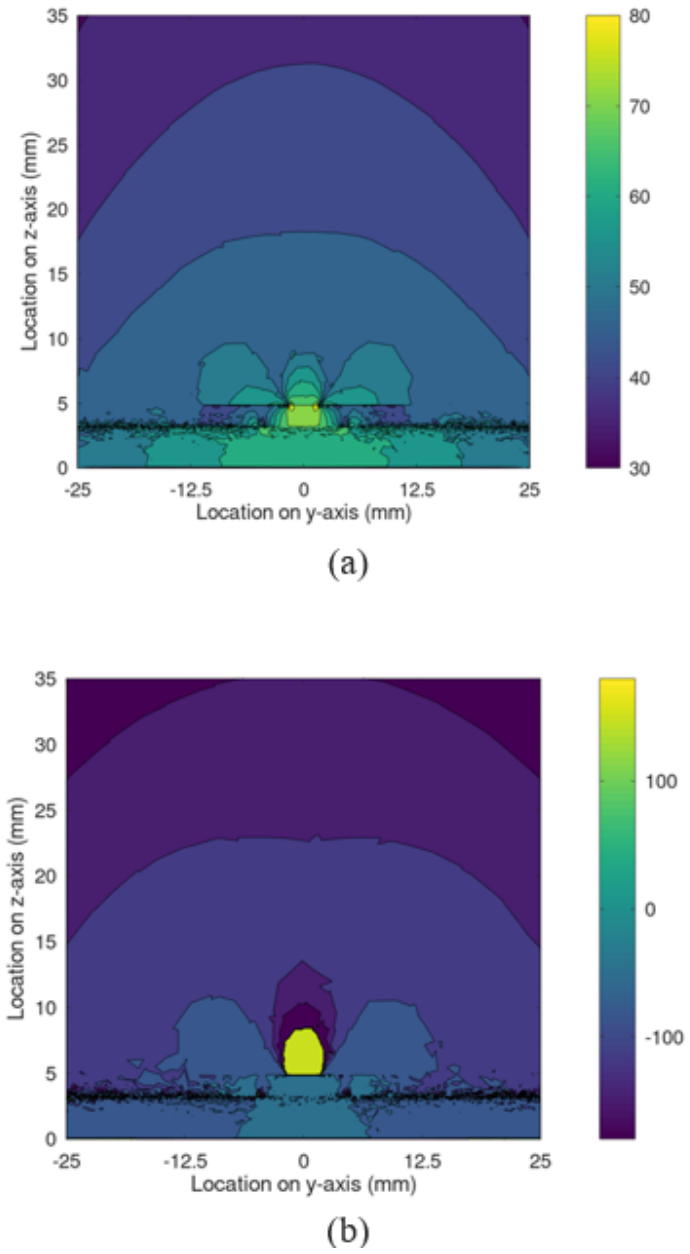


Figure 9

Near field performance of the antenna in absence of superstrate (a) E-field magnitude on a plane, (b) E-field phase on a plane.

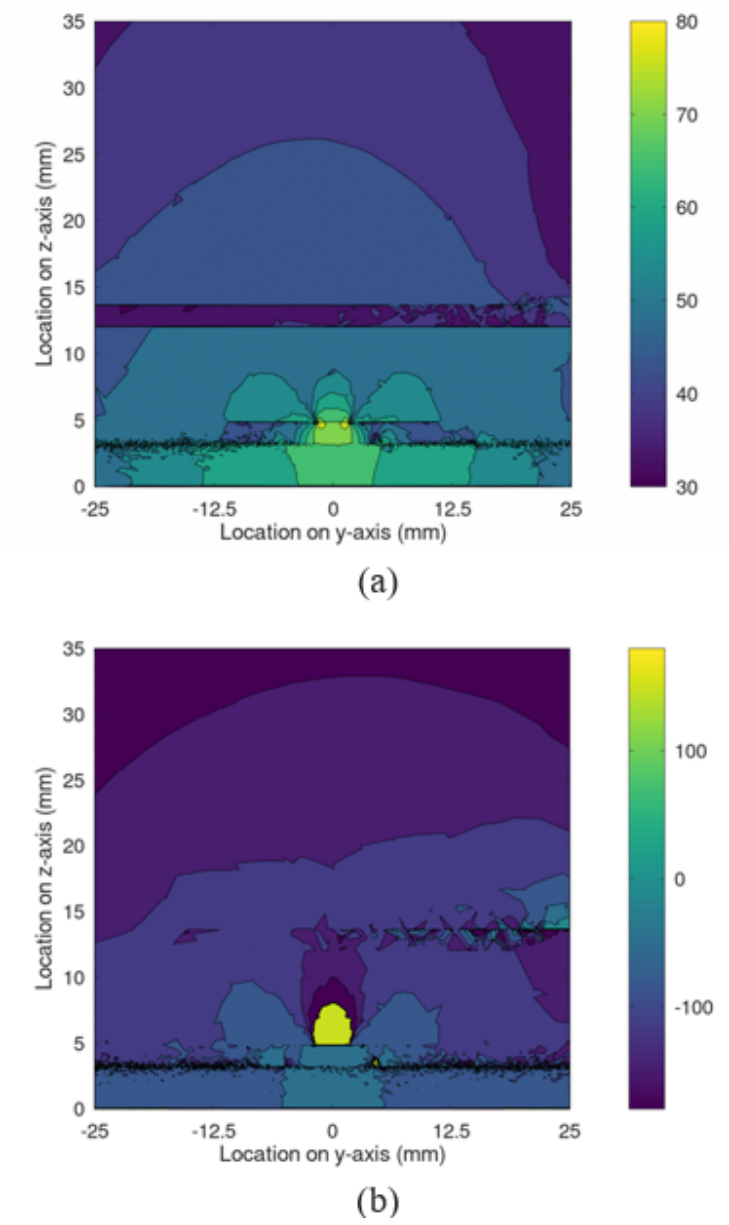


Figure 10

Near field performance of the antenna with superstrate (a) E-field magnitude on a plane, (b) E-field phase on a plane

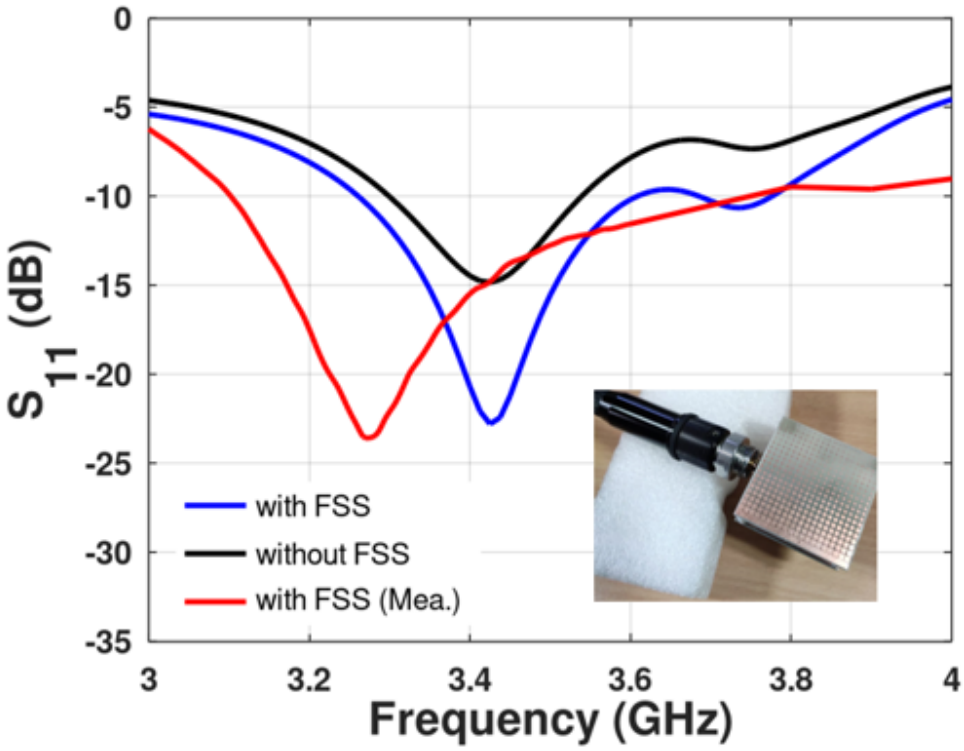


Figure 11

Measured and simulated S_11of the proposed antenna.

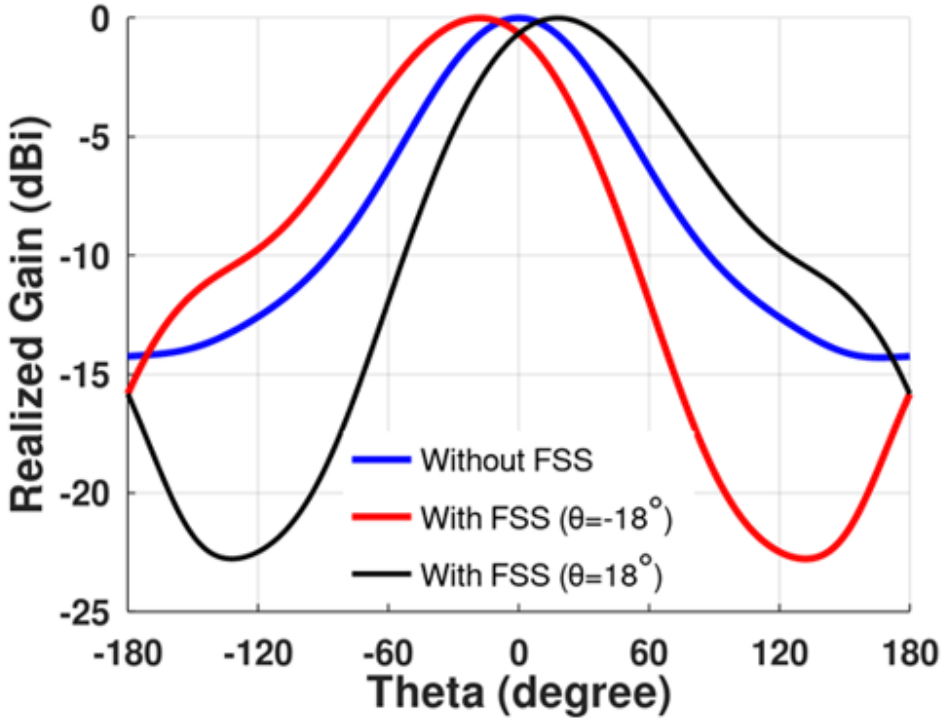


Figure 12

Radiation pattern in absence and presence of superstrate.

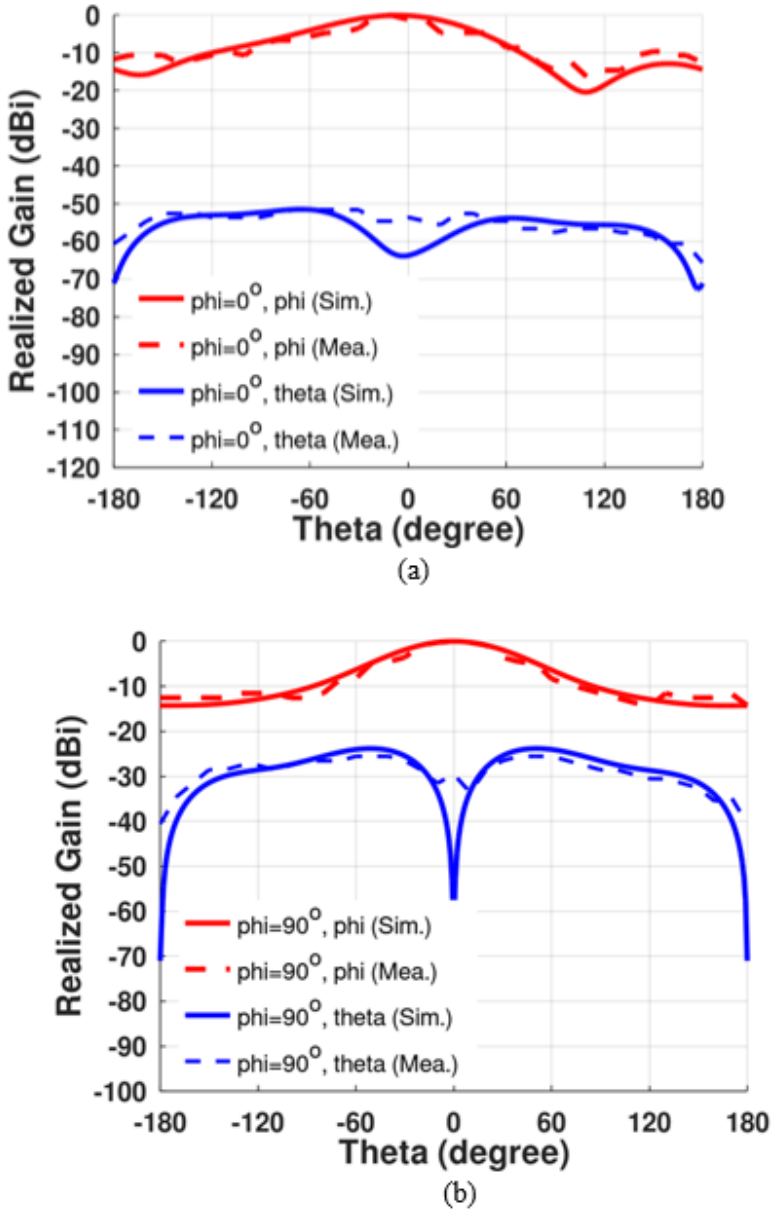


Figure 13

Simulated and measured radiation pattern of the antenna without superstrate, (a) φ=0° cut (b) φ=90° cut.

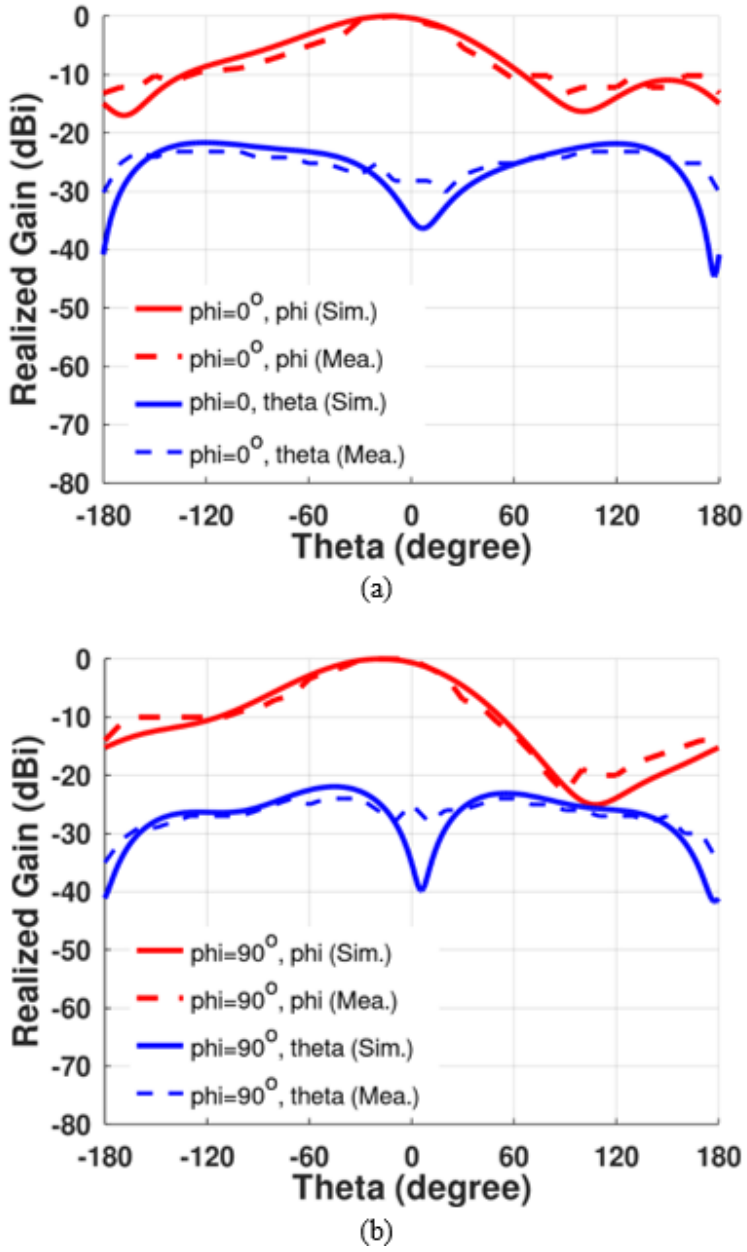


Figure 14

Simulated and measured radiation pattern of the antenna with superstrate, (a) φ=0° cut (b) φ=90° cut.

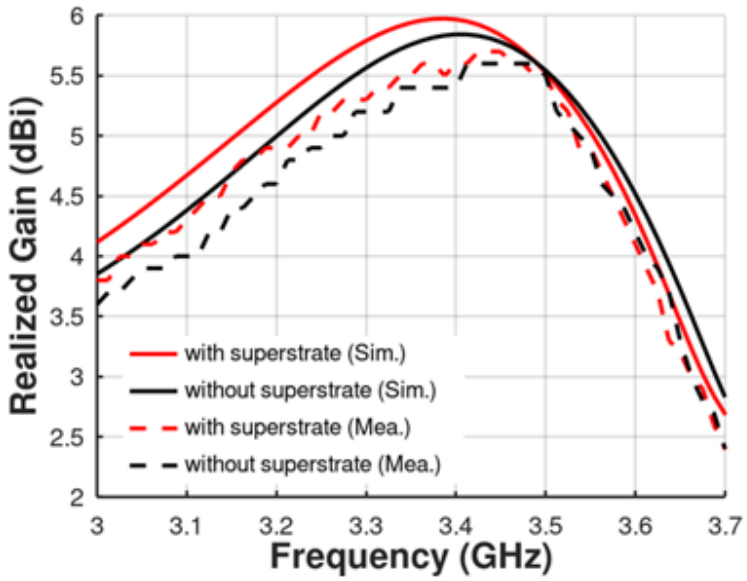


Figure 15

Peak realized gain of the antenna when the beam is steered in different directions.

2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
September 27 - October 1, 2021. Prague, Czech Republic

2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) | 978-1-6654-1714-3/21/$31.00 ©2021 IEEE | DOI: 10.1109/IROS51168.2021.9635841

Modeling and Control of PANTHERA Self-Reconfigurable Pavement Sweeping Robot under Actuator Constraints

Madan Mohan Rayguru[1], M. R. Elara[1], A. A. Hayat[1], B. Ramalingam[1] and Spandan Roy[2]

***Abstract*— The focus of this paper is (i) to derive a suitable dynamic model for a self-reconfigurable pavement sweeping robot PANTHERA and (ii) to design a robust controller for the same to tackle uncertainties stemming from the reconfiguration process, external disturbances and from actuator saturation. To meet the first objective, an Euler-Lagrangian dynamic model is proposed to incorporate the effects of configuration changes on the system dynamics. Based on this model, the second objective is met via designing a singular perturbation based robust controller which can tackle the aforementioned uncertainties without violating the actuation limits. To circumvent the vulnerability toward actuator saturation, the proposed controller is built on contraction theory, which, compared to a conventional Lyapunov theory based design, allows to improve closed-loop tracking performance without reducing the singular perturbation parameter. Experimental results on the PANTHERA reconfigurable robot validate the effectiveness of the proposed controller over the state of the art.**

I. Introduction

The development of urban pavement infrastructure systems is an integral part of modern city expansion processes. However, maintenance of large infrastructure is quite labour intensive and involves health risks. Therefore, multiple automatic/semi-automatic solutions have been developed to circumvent the issues involved in the conventional manual process [1]–[5]. Nevertheless, these fixed-structured based platforms are not suitable when the width of pavement changes (e.g., when pavement becomes narrow, a wider robotic platform cannot be used). Further, it is not feasible either operationally or economically to use different cleaning robots depending on the width of a pavement.

Owing to the inherent capability to autonomously change the shapes/configurations, re-configurable robots are becoming a viable alternative for fixed-structured robots while operating in a dynamics environment mentioned earlier [2], [6]–[9]. In view of this advantage of a re-configurable mechanism, the pavement cleaning robot PANTHERA was developed in [9], [10]. Henceforth, the term 'self-reconfigurable robot' will be denoted as 'reconfigurable robot' for brevity.

This research is supported by the National Robotics Programme under its Robotics Enabling Capabilities and Technologies (Funding Agency Project No. 192 25 00051), National Robotics Programme under its Robot Domain Specific (Funding Agency Project No. 192 22 00058) and administered by the Agency for Science, Technology and Research.

[1] Madan Mohan Rayguru is with the department of Electrical Engineering, Delhi Technological University, Delhi, India (mmrayguru87@gmail.com)

[2] M. R. Elara, A. A. Hayat and B. Ramalingam are with Engineering Product Development Pillar in Singapore University of Technology and Design, Singapore (rajeshelara@sutd.edu.sg)

[3] S. Roy is with Robotics Research Center, International Institute of Information Technology Hyderabad (IIIT-H), Hyderabad, India (spandan.roy@iiit.ac.in)

With increasing demand in reconfigurable mobile robots, the modeling and control aspects of such systems have to be deliberated for ensuring better performance. A few notable works regarding control of reconfigurable mobile robots include PID [6], [10], backstepping [11], instantaneous rotation method (ICR) [2], resolved motion rate method [12], [13], fuzzy logic technique [14] and numerical optimization approach [15]. However, these state-of-the-art designs rely on the kinematic model of the robot: such model is not suitable to handle dynamic uncertainties such as perturbation in inertia, friction forces, external disturbances etc. For example, the configuration/shape changes in PANTHERA lead to variations in wheel width, which in turn affects dynamical properties like moment of inertia and center of mass (COM) [16], [17].

In view of the above discussion, a control scheme to suitably tackle dynamical perturbations in a reconfigurable mobile robot like PANTHERA, stemming from its reconfiguration process and from unknown disturbances, to the best of the knowledge of the authors, is still missing at large. In addition, one has to inevitably consider the limited actuation capabilities in real-life scenarios. Toward this direction, this work proposes a contraction theory based robust controller for PANTHERA with the following notable features:

- Unlike the existing kinematic model based designs, a suitable dynamic model of PANTHERA is derived by considering the effects of reconfiguration process and unknown disturbances.
- Compared to the conventional Lyapunov based designs, the proposed contraction theory based design can circumvent the vulnerability of actuator saturation while improving closed-loop tracking performance (cf. Remark 2 for details).

The effectiveness of the proposed design is extensively verified via experiments on PANTHERA robot in comparison with the state of the art. Though primarily developed for PANTHERA, the proposed modelling and control scheme is developed in a quite generalized framework and can be applied to other similar reconfigurable mobile robots.

The following notations are used in this paper: $||.||$ denotes Eucledian norm, $\tanh(.)$ denotes Tan-hyperbolic function, and any positive definite matrix $(\cdot)$ is defined as $(\cdot) > 0$.

The remainder of the paper is organised as follows: Sect. II presents the dynamic model of PANTHERA; Sect. III describes the proposed controller and the closed-loop stability analysis; Sect. IV provides comparative experimental results, while concluding comments are provided in Sect. V.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 06:18:28 UTC from IEEE Xplore. Restrictions apply.

TABLE I

PHYSICAL PARAMETERS OF PANTHERA

Parameter	Magnitude
Robot Width (b)	Compressed: 0.5m Expanded: 1.1m
Robot length ($2d$)	1.2m
Wheel radius (r_w)	0.1m
Platform mass (m)	200kg
Wheel inertia (I_w)	$0.8kgm^2$
Platform Inertia ($\bar{I}$)	Compressed: $28.3kgm^2$ Expanded: $44.2kgm^2$

II. DYNAMIC MODELING OF PANTHERA

PANTHERA follows a differential drive mechanism, where its capability to change configuration (c.f. Fig. 1) stems from a lead-screw motor (c.f. Fig. 2). A detailed mechanical architecture of the PANTHERA platform is described in [10], [18], while its physical parameters are summarized in Table I. The Euler-Lagrange (EL) equations of motion for PANTHERA, by accounting for the unmodelled dynamics components arising from wheel friction, external disturbances etc. can be found from the methods described in [19], [20] as

Fig. 1. PANTHERA platform in compressed and expanded configurations.

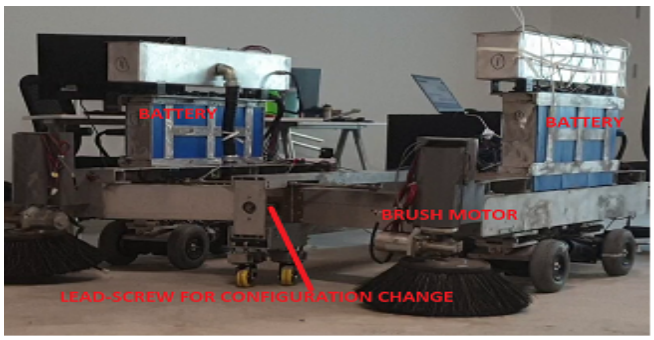


Fig. 2. Architecture of PANTHERA base.

$$M_R\ddot{q}_R + f(\dot{q}_R) + d_s = \tau, \quad \tau = [\tau_r \ \tau_l]^T, \tag{1}$$

$$\dot{q} = \underbrace{\begin{bmatrix} \frac{r_w}{b}\left(\frac{b}{2}\cos(\varphi) - d\sin(\varphi)\right) & \frac{r_w}{b}\left(\frac{b}{2}\cos(\varphi) + d\sin(\varphi)\right) \\ \frac{r_w}{b}\left(\frac{b}{2}\sin(\varphi) + d\cos(\varphi)\right) & \frac{r_w}{b}\left(\frac{b}{2}\sin(\varphi) - d\cos(\varphi)\right) \\ r_w/b & -r_w/b \\ 1 & 0 \\ 0 & 1 \end{bmatrix}}_{S(q)} \dot{q}_R, \tag{2}$$

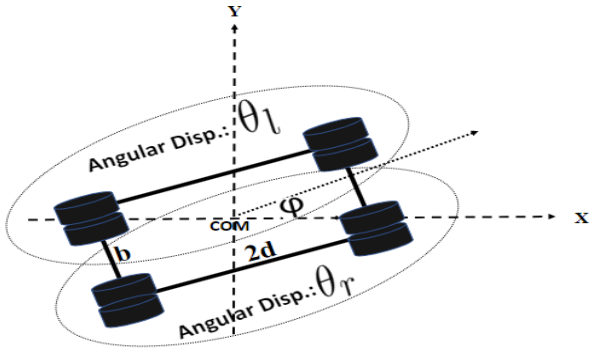


Fig. 3. Schematic of the base of PANTHERA.

$$q_R = [\theta_r \ \theta_l]^T, M_R = S^T M S = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}, \tag{3}$$

$$M_{11} = M_{22} = I_w + \{\bar{I} + m(b^2/4 - d^2)\}(r_w{}^2/b^2),$$
$$M_{12} = M_{21} = \{m(b^2/4 + d^2) - \bar{I}\}(r_w{}^2/b^2) \tag{4}$$

where the state vector $q \in \mathbb{R}^5 = \{x_c, y_c, \varphi, \theta_r, \theta_l\}$ comprises of the position of center of mass (COM), orientation of the robot, and angular displacements of the wheels from two sides; the variables (τ_l, τ_r) denote the wheel torques from left and right sides respectively; m denotes mass of the platform; $\bar{I}$ indicates the platform inertia; I_w denotes inertia of the wheels; the unmodelled friction terms are captured in $f(\dot{q}_R)$; other system parameters are described in Table I. The dynamic model (1) can be further simplified as:

$$\begin{aligned} \dot{q}_{11} &= q_{12} \\ \dot{q}_{12} &= \bar{M}_{11}\tau_l + \bar{M}_{12}\tau_r + f_1 + d_1 \\ \dot{q}_{21} &= q_{22} \\ \dot{q}_{22} &= \bar{M}_{21}\tau_l + \bar{M}_{22}\tau_r + f_2 + d_2 \end{aligned} \tag{5}$$

where $q_{11} = \theta_r$, $q_{12} = \dot{\theta}_r$, $q_{21} = \theta_l$, $q_{22} = \dot{\theta}_l$, $\bar{M}_{ij} = M_R^{-1}(i,j)$, $i,j \in \{1,2\}$, $M_R^{-1}d_s = [d_1, d_2]^T$ and $M_R^{-1}f = [f_1, f_2]^T$.

Impact of Reconfiguration on the System Dynamics: The inertia terms M_{ij} in (4) and disturbance terms d_i, f_i $i = 1, 2$ in (5) vary as the distance between wheels and the COM of the platform is changed during reconfigurations. This phenomenon can be captured in the dynamics as

$$\begin{aligned} \dot{q}_{11} &= q_{12} \\ \dot{q}_{12} &= (\bar{M}_{11} + \Delta\bar{M}_{11})\tau_l + (\bar{M}_{12} + \Delta\bar{M}_{12})\tau_r + \delta d_r \\ \dot{q}_{21} &= q_{22} \\ \dot{q}_{22} &= (\bar{M}_{21} + \Delta\bar{M}_{21})\tau_l + (\bar{M}_{22} + \Delta\bar{M}_{22})\tau_r + \delta d_l \end{aligned} \tag{6}$$

where $\Delta\bar{M}_{ij}$ stems from the variations in M_R and

$$[\delta d_r, \delta d_l]^T = (M_R + \Delta M_R)^{-1}d_s + (M_R + \Delta M_R)^{-1}f.$$

Further, input torques are also bounded due to the imposed actuator limits. Hence, the terms $\Delta\bar{M}_{ij}\tau_i$ in (6) can be merged with $\delta d_l, \delta d_r$, and the *PNATHERA dynamics under*

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 06:18:28 UTC from IEEE Xplore. Restrictions apply.

reconfigurations can be written in a compact form as:

$$\begin{aligned}\dot{q}_{11} &= q_{12}\\ \dot{q}_{12} &= \bar{M}_{11}\tau_l + \bar{M}_{12}\tau_r + \delta d_1\\ \dot{q}_{21} &= q_{22}\\ \dot{q}_{22} &= \bar{M}_{21}\tau_l + \bar{M}_{22}\tau_r + \delta d_2.\end{aligned} \tag{7}$$

where $\delta d_1 = \delta d_r + \Delta\bar{M}_{11}\tau_l + \bar{M}_{12}\tau_r, \delta d_2 = \delta d_l + \Delta\bar{M}_{21}\tau_l + \bar{M}_{22}\tau_r$. The following standard assumption is considered:

Assumption 1: The disturbances $\delta d_1, \delta d_2$ satisfy

$$||\delta d_1|| \leq c_{lo}, ||\delta d_2|| \leq c_{ro}$$

where c_{lo}, c_{ro} are known positive constants.

Remark 1: Assumption 1 is quite common in literature and it is satisfied for slowly varying uncertainties [21]–[23]. The pavement sweeping task is generally slow and PANTHERA does not move during the reconfiguration process (static reconfiguration) to avoid collision in narrow pavements. Therefore, Assumption 1 holds true for this platform.

Control Goal: Under Assumption 1, to develop a robust controller for the PANTHERA robot such that its trajectories from (7) follow some desired path generated from the following stable linear dynamical system

$$\dot{x}_{dj} = A_{dj}x_{dj} + B_{dj}r_j,\ j = 1, 2$$
$$A_{dj} = \begin{bmatrix} 0 & 1 \\ -a_{dj1} & -a_{dj2}\end{bmatrix},\ B_{dj} = \begin{bmatrix} 0 \\ b_{dj}\end{bmatrix}, \tag{8}$$

where $x_{dj} = [x_{dj1}, x_{dj2}]^T \in R^2$ are the state vectors for the desired subsystems $[q_{j1}, q_{j2}]^T$ and r_j is a smooth feedforward signal. The matrix elements a_{dj1}, a_{dj2} should be designed to make A_{dj} a Hurwitz matrix and the feedforward signal r_j can be utilized to generate the desired paths.

III. Controller Design and Analysis

This section proposes a contraction theory based dynamic high gain controller for the PANTHERA dynamics (7) to address the actuator constraints via timescale separation and singular perturbation. Before presenting the proposed controller, a few design preliminaries are provided:

A. Preliminaries: Contraction and Partial Contraction

A nonlinear smooth system defined as

$$\dot{x} = f(x). \tag{9}$$

is *contracting*, if any pair of arbitrary trajectories $(x_1(t), x_2(t))$ of this system converge towards each other irrespective of their starting points $(x_1(0), x_2(0))$ [24]–[26]. A sufficient condition for contraction of (9) is given by

$$\Big((\dot{\Theta} + \Theta\frac{\partial f}{\partial x})\Theta^{-1}\Big) + \Big((\dot{\Theta} + \Theta\frac{\partial f}{\partial x})\Theta^{-1}\Big)^T \leq -\lambda I, \tag{10}$$

where Θ denotes a positive definite matrix and $\lambda \in R^+$. When the condition (10) holds true, we define the system to be contracting with (λ, Θ), with Θ being the associated transformation matrix, and λ being the contraction rate.

The idea of contraction can be more generalized via using the concept of *partial contraction* [27], [28]. Consider a smooth nonlinear system

$$\dot{x} = f(x, y, t) \tag{11}$$

with a virtual/auxiliary system $\dot{z} = f(z, y, t)$ (cf. [27]). If the virtual system is contracting in z for any value of y, $\forall t > 0$, then (11) is defined to be *partially contracting* in x. The virtual and the actual system posses similar mathematical model, and by the virtue of partial contraction, they follow same similar stare trajectories [7], [27].

B. Proposed Controller Design and Analysis

The controller under actuator constraints is designed as:

$$\begin{aligned}&\tau_l = \sigma(v_l),\ \tau_r = \sigma(v_r),\\ &\mu\dot{v} = \mu\begin{bmatrix}\dot{v}_l \\ \dot{v}_r\end{bmatrix} = -F(q, v, x_d, r),\end{aligned} \tag{12}$$

$$F(.) = \begin{bmatrix}\bar{M}_{11}\sigma(v_l) + \bar{M}_{12}\sigma(v_r) + \sum_{i=1}^{2}(a_{d1i}x_{d1i} + a_{e1i}e_{1i}) \\ -b_{d1}r_1 \\ \bar{M}_{21}\sigma(v_l) + \bar{M}_{22}\sigma(v_r) + \sum_{i=1}^{2}(a_{d2i}x_{d2i} + a_{e2i}e_{2i}) \\ -b_{d2}r_2\end{bmatrix}$$

where $\sigma(.) = \beta\tanh((.)/\beta)$ with β being the actuator limit; $\mu \in (0, 1)$ is the singular perturbation parameter; $q = [q_{11}, q_{12}, q_{21}, q_{22}]^T$; $e_{1i} = q_{1i} - x_{d1i}, e_{2i} = q_{2i} - x_{d2i}$ $i = 1, 2$; a_{e1i}, a_{e2i} are positive design scalars. Small value of μ makes the control variables v_l, v_r evolve faster and the closed-loop system can be considered as a two timescale singularly perturbed system.

Let us define a Hurwitz matrix

$$A_{ej} = \begin{bmatrix} 0 & 1 \\ -a_{e1j} & -a_{e2j}\end{bmatrix} \tag{13}$$

with distinct eigenvalues and a_{e1j}, a_{e2j} $j = 1, 2$ are positive design scalars. Therefore, there exists an associated positive definite transformation matrix Θ_{Ae} (can be formed by the eigenvectors of A_{ej}) such that $\bar{A}_e = \Theta_{Ae}A_{ej}\Theta_{Ae}^{-1}$ is negative definite.

Let λ_{ae} be the magnitude of the minimum eigenvalue (i.e. $-\lambda_{ae}$) of A_{ej}, which also serves the purpose of contraction rate for the error dynamics (cf. the discussion below (23)); the left wheel tracking error variables be $e_1 = [e_{11}, e_{12}]$ and the right wheel tracking error variables be $e_2 = [e_{21}, e_{22}]$. The closed-loop stability result is stated below:

Theorem 1: Under Assumption 1, the tracking error for the system (7) with the control law (12) remains bounded as

$$\limsup_{t\to\infty} ||e_1(t)|| \leq \frac{\chi_{ae}L_1}{\lambda_{ae}}\Big(\mu\frac{c_2}{\lambda_v} + c_{lo}\Big), \tag{14a}$$

$$\limsup_{t\to\infty} ||e_2(t)|| \leq \frac{\chi_{ae}L_2}{\lambda_{ae}}\Big(\mu\frac{c_2}{\lambda_v} + c_{ro}\Big), \tag{14b}$$

where L_1, L_2 are the Lipschitz constants corresponding to the left and right wheel dynamics of $F(\cdot)$ (cf. the discussion below (32)) respectively; χ_{ae} is the condition number of Θ_{Ae}; c_2 and λ_v are constants (cf. (27)).

Proof: See Appendix. ■

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 06:18:28 UTC from IEEE Xplore. Restrictions apply.

Remark 2 (Lyapunov vs. Contraction theory based design): Conventional Lyapunov-based designs require μ to be very small for satisfactory closed-loop performance, which may lead to actuator saturation due to excessive control action. Whereas, for the proposed design, (14) reveals that, beside μ, the error bounds also depend on contraction rate λ_{ae} and condition number χ_{ae}. Therefore, the tracking accuracy can be tuned via λ_{ae}, χ_{ae}, without decreasing μ. As a result, vulnerability toward actuator saturation can be circumvented by the proposed contraction theory based design. On the other hand, the effect of reconfiguration is reflected in the tracking error through the terms c_{lo}, c_{ro}.

Remark 3 (On the selection of design parameters): By adequately selecting the design scalars a_{eij} of the matrix A_{ej} in (13), one can tune λ_{ae} and χ_{ae} to achieve desired control performance via (14).

IV. Experimental Results and Discussion

The PANTHERA platform is run on eight wheels, where each pair of wheels are powered through a single motor driver (c.f. Fig. 4). By following the differential drive technique, the same side wheels receive same control input (τ_l for motor drivers 1, 2 and τ_r for motor drivers 3, 4). The scrubber motors (motor driver 6) for sweeping is not used for the experiments, and motor driver 5 (for lead-screw enabled configuration change) is kept independent from the control scheme. The associated sensor suits and data acquisition process for the platform is described in [20].

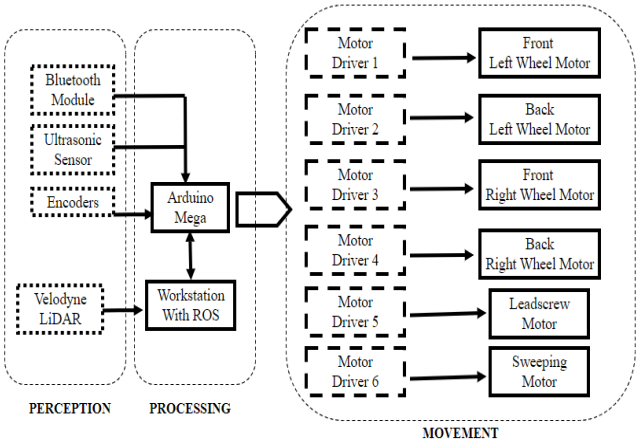


Fig. 4. Schematic of electronics module in PANTHERA.

A. Experimental Scenarios and Results

The experiments are designed and performed to validate the potential of the proposed controller (12) for the PANTHERA platform modeled as (7). A $2m$ radius desired circular path is designed for the robot to track, by choosing two separate velocities for the wheels from each side [19]. This desired circular trajectory can be realized through setting the desired wheel positions as ($\theta_r = q_{11}, \theta_l = q_{21}$)

$$x_{d11} = \theta_r^d = 4t \ rad, \quad x_{d21} = \theta_l^d = 3.3t \ rad,$$

which is generated via the reference model (8) with

$$A_{d1} = A_{d2} = \begin{bmatrix} 0 & 1 \\ -20 & -21 \end{bmatrix}, B_{d1} = B_{d2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (15)$$
$$r_1 = 40t, r_2 = 33t.$$

The desired (x_c, y_c) positions for the PANTHERA platform is derived via (3). Two different experimental scenarios are chosen as follows:

1) Scenario 1: In this scenario, the tracking performance of PANTHERA is analysed by tuning the contraction based design parameters λ_{ae}, χ_{ae}, while without changing the singular perturbation parameter μ. Two cases of design parameter selection are considered for the proposed controller (12): (i) case 1: $a_{e11} = a_{e21} = 10$, and $a_{e12} = a_{e22} = 11$; (ii) case 2: $a_{e11} = a_{e21} = 12$, and $a_{e12} = a_{e22} = 7$. For these choices of controller parameters, the contraction rates and condition number becomes $\lambda_{ae} = 1$ for case 1, and $\lambda_{ae} = 3$ for case 2; $\chi_{ae} = 22$ for case 1 and $\chi_{ae} = 16$ for case 2. Other parameters are fixed at $\mu = 0.1$, and the actuation limit at $\beta = 5$. All the experiments are performed with the same initial position of the PANTHERA, i.e. $x_c(0) = 9.5, y_c(0) = -2.5$ (from LiDAR mapping). Two configurations of the PANTHERA are selected for experimental evaluation: (i) Configuration 1: fully stretched and (ii) Configuration 2: fully compressed. The change form configuration 1 to 2 is carried out at $t = 45$ sec.

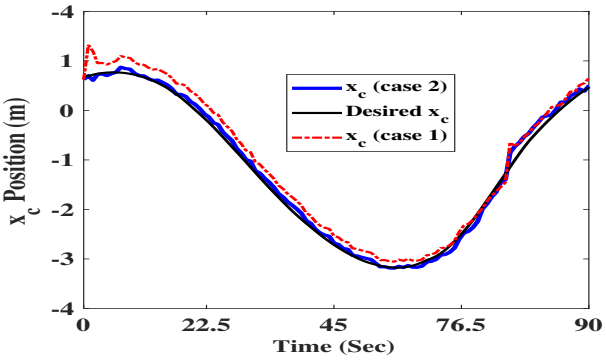


Fig. 5. Tracking in x_c position for various cases.

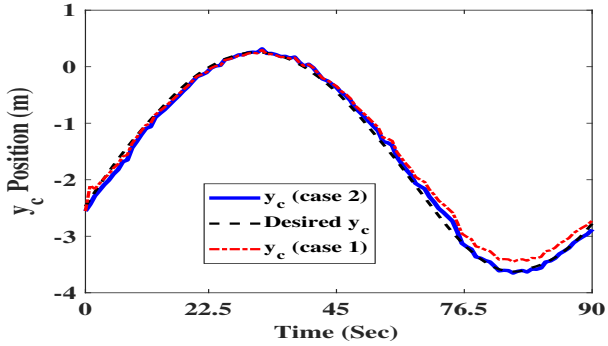


Fig. 6. Tracking in y_c position for various cases.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 06:18:28 UTC from IEEE Xplore. Restrictions apply.

The Figs. 5, 6 present the closed loop tracking performance for cases 1 and 2. These plots reveal the superiority of the controller performance for case 2 against case 1. Compared to case 1, the choice of design parameters for case 2 leads to larger contraction rate λ_{ae}, and smaller condition number χ_{ae}. Therefore, in line with Remark 2, improved tracking accuracy is observed for case 2 which also confirms the effectiveness of the contraction theory based design.

2) Scenario 2: This scenario examines the performance of the proposed controller against a kinematics based PID control [6], and the sliding mode control (SMC) [29]. In this scenario, the design parameters for proposed controller are chosen same as the case 2 of Scenario 1. The sliding surface and switching gain for SMC is chosen as, $S = 3e_i + 2\int e_i(t) + \dot{e}_i$, and $k_s = 100$ respectively. The comparative performances of these controllers are presented in Table II in terms of root-mean-squared (RMS) path error (Euclidean distance in x_c and y_c position error), where, the performance improvements of the proposed controller against the respective controllers are shown as percentage values.

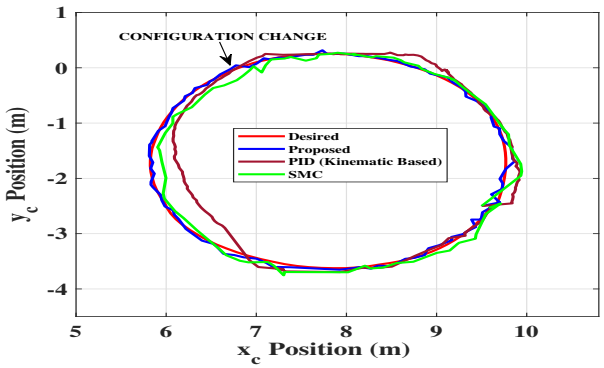


Fig. 7. Circular path tracking comparison in Scenario 2.

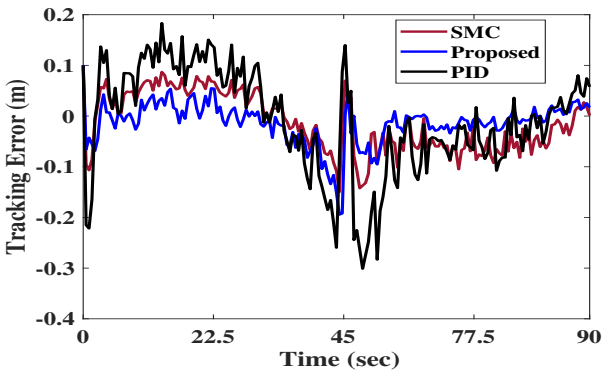


Fig. 8. Path tracking error comparison for Scenario 2.

The results shown in Figs. 7-8 and Table II highlight that the tracking accuracy of PID controller is low and suffers from transient overshoots arsing from configuration change as the kinematic model excludes external and frictional forces. The performance of SMC suffers because it does not account for the uncertainty stemming from configuration changes as well as actuator saturation, leading to larger performance drop after the reconfiguration (for $t > 45$s) in Fig. 8. Compared with these controllers, the proposed technique ensures better performance, as it incorporates actuator saturation in the design, and the contraction based design parameters are judiciously tuned. Note that, the control inputs always remain within the prescribed limit (Fig. 9), even though the inputs saturates for very short time to compensate the effects of reconfiguration.

TABLE II
RMS PATH ERROR (M) COMPARISON

Controller	Configuration 1	Configuration 2
PID	0.009 (53%)	0.015 (70%)
SMC	0.008 (25%)	0.012 (40%)
Proposed	0.006	0.008

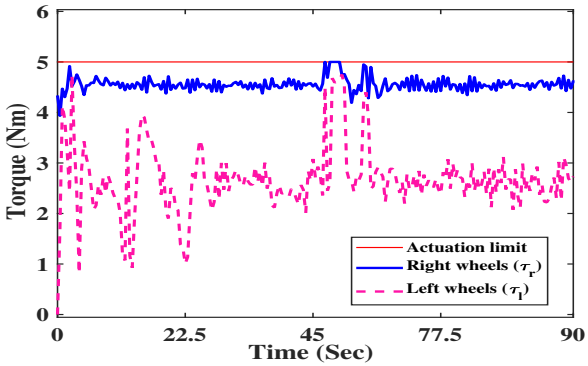


Fig. 9. Control inputs from the proposed controller

V. CONCLUSION

This work proposed an Euler-Lagrangian method based dynamic model and a control framework for PANTHERA self-reconfigurable mobile robot. A dynamic controller based on timescale separation was advocated to address actuator limits. Further, the closed-loop system stability was investigated via contraction theory: this, in contrast to a conventional Lyapunov based method, allowed to tune closed-loop tracking performance without reducing singular perturbation parameter and to reduce vulnerability toward actuator saturation. Extensive experimental results revealed the effectiveness of proposed methodology.

APPENDIX

Proof of Theorem 1: The closed-loop stability analysis rests on three stages: convergence (A) of trajectories for reduced order slow system, (B) of dynamic control inputs (v_l, v_r) and (C) of tracking error. We present Lemma 1, which is later used for the stability analysis.

Lemma 1 ([7], [30]): Consider the perturbed nonlinear system (9) to be contracting with (λ, Θ) as

$$\dot{x}_p = f(x_p) + d(x_p, \Delta), \tag{16}$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 06:18:28 UTC from IEEE Xplore. Restrictions apply.

where the term Δ denotes the additive uncertainties in the system. The following arguments hold true:

(i) If there exists a known positive scalar d_o such that $||d(x_p, \Delta)|| \leq d_o$, then the following holds:

$$||x_p(t) - x(t)|| \leq \Upsilon \exp^{-\lambda t} ||x_p(0) - x(0)|| + (\Upsilon d_o)/\lambda, \quad (17)$$

where Υ stand for the supremum of the condition number of the associated transformation matrix Θ.

(ii) Consider the case where $||d(x_p, \Delta)|| \leq d_{on} + d_{1n}||x_p||$, $d_{on}, d_{1n} \in R^+$. If the nominal system (9) has a forward bounded solution x_{oo}, and $d_{1n} \leq \frac{\lambda}{\Upsilon}$, then

$$||x_p(t) - x(t)|| \leq \Upsilon \exp^{-(\lambda - \Upsilon d_{1n})t} ||x_p(0) - x(0)|| + d_{2n} \quad (18)$$

where $d_{2n} = \frac{\Upsilon(d_{on} + d_{1n} x_{oo})}{\lambda - \Upsilon d_{1n}}$.

Subsequently, we discuss various stages of the stability analysis.

A. Convergence of Reduced Slow System

The Jacobian J_c for the controller dynamics (12) w.r.t (v_l, v_r) can be expressed as:

$$J_c = -\frac{1}{\mu} M_R^{-1} \begin{bmatrix} \frac{\partial \sigma(v_l)}{\partial v_l} & 0 \\ 0 & \frac{\partial \sigma(v_r)}{\partial v_r} \end{bmatrix}, \quad (19)$$

where $M_R^{-1} > 0$ by system property. Further, $\frac{\partial \sigma(v_i)}{\partial v_i} > 0$, $(v_i = v_l$ or $v_r)$ in any bounded and compact region owing to the properties $\tanh(\cdot)$. So, J_c is uniformly negative definite and a positive constant $\frac{\bar{\lambda}_v}{\mu}$ exists such that the inequality $J_c + J_c^T \leq \frac{\bar{\lambda}_v}{\mu} I_2$ is satisfied. Then, from the definition of partial contraction [27], the closed-loop singularly perturbed dynamics is partially contracting in the control variables $v = [v_l, v_r]^T$.

The vector field $F(.)$ of (12) constitutes a partially contracting dynamics (w.r.t v), and therefore, $F(.) \to 0$ results in a unique mapping between v and other arguments [27], [28]. In other words, the vector field $F(.)$ has a unique root v^* such that $F(v^*, q, x_{dj}, r_j) = 0$, i.e.,

$$F(.) \to 0 \Rightarrow v \to v^* = h_c(q_{ij}, x_{dij}, r_j). \quad (20)$$

where $h_c(.) = V^*$ is the unique solution of the equation $F(.) = 0$. In singular perturbation literature h_c is defined as *slow manifold*.

The reduced slow system is obtained when the slow manifold $h_c(.)$ replaces v in the PANTHERA dynamics. From the structure of $F(.)$, $F(.) \to 0$ results in

$$\begin{bmatrix} \bar{M}_{11}\sigma(v_l) + \bar{M}_{12}\sigma(v_r) \\ \bar{M}_{21}\sigma(v_l) + \bar{M}_{22}\sigma(v_r) \end{bmatrix} \to \begin{bmatrix} \sum_{i=1}^{2}(a_{d1i}x_{d1i} + a_{e1i}e_{1i}) \\ -b_{d1}r_1 \\ \cdots\cdots\cdots \\ \sum_{i=1}^{2}(a_{d2i}x_{d2i} + a_{e2i}e_{2i}) \\ -b_{d2}r_2 \end{bmatrix} \quad (21)$$

The error dynamics can be written as:

$$\dot{e}_j = A^o x_j + BF_j(v, q, x_d, r) + \delta d_j - A_{dj}x_{dj} - B_{dj}r_j \quad (22)$$

where $e_j = [e_{j1}, e_{j2}]$, $e_{ji} = q_{ji} - x_{dji}$, $i, j = 1, 2$, $F_j(.)$ denotes the jth element of $F(.)$, and

$$B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, A^0 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

From (21) and (22), it can be derived that the reduced order error dynamics is derived to be

$$\dot{e}_j = A_{ej}e_j + \delta d_j, A_{ej} = \begin{bmatrix} 0 & 1 \\ -a_{e1j} & -a_{e2j} \end{bmatrix}. \quad (23)$$

From the discussion below (13), the dynamics (23) can be treated as a perturbed contracting system with metric Θ_{Ae}.

B. Convergence of Control Variables

Any controller should ensure boundedness of all internal variables. The contracting error dynamics in the previous section was obtained for the ideal scenario where $v \to h_c(.)$, which may not be satisfied if μ in (12) is very small. To make the analysis more practical, define an error vector $v_e = v - h_c(.)$. The dynamics of v_e in a stretched time scale (fast boundary layer dynamics) can be derived as

$$\frac{dv_e}{ds} = F(q, v_e + h_c(.), x_d, r) - \mu \dot{h}_c(.). \quad (24)$$

where $s = \frac{t}{\mu}$ denotes the stretched time scale. Without considering the perturbation term in (24), the nominal boundary layer dynamics can be written as

$$\frac{dv_e}{ds} = F(q, v_e + h_c(.), x_d, r). \quad (25)$$

It is already proven that the Jacobian of $F(.)$ w.r.t v (19) is negative definite. So, the partial contraction of (25) w.r.t v_e is guaranteed, where the contraction rate is $(\bar{\lambda}_v/\mu)$ and transformation matrix $\Theta_v = I$ (condition number $\chi_v = 1$). From the smoothness properties of $h_c(.)$ and its derivative, one can derive the following for some $c_{2a}, c_{2b} \in \mathbb{R}^+$

$$||\dot{h}_c(.)|| \leq c_{2a}||v_e|| + c_{2b}.$$

For any initial conditions inside a compact region $(D_q \times D_v \times D_{\hat{q}})$, and any $0 < \mu < 1$ which satisfies

$$c_{2a} \leq \bar{\lambda}_v/\mu^2 \quad (26)$$

the boundary layer dynamics (24) is assured to be partially contracting in v_e by the virtue of the Lemma 1. Further, there exists a forward bounded solution v_o ($||v_{ds}(s)|| \leq v_o \forall\, t \geq 0$) for the unperturbed partially contracting system (25). Using the Lemma 1, we can derive

$$\limsup_{t \to \infty} ||v_e(t)|| = ||v(t) - h_c(.)|| \leq \mu \frac{c_2}{\lambda_v}, \quad (27)$$

where $\lambda_v = \bar{\lambda}_v - \mu c_{2a}$ and $c_2 = \mu(c_{2a}v_o + c_{2b})$.

C. Convergence of Tracking Errror

The reduced slow error dynamics was proved to be a perturbed contracting system in Section IV.A. In the absence of any uncertainties ($\delta d_j = 0$) in (21),

$$e_j \to 0 \Rightarrow q_{ij} \to x_{dij},\ i, j = 1, 2. \quad (28)$$

For deriving the bounds for tracking errors, express the closed-loop PANTHERA dynamics (7) as:

$$\begin{aligned} \dot{q}_j^l &= A^o q_j^l + BF_j^l(v, \delta d_1) \\ \dot{q}_j^r &= A^o q_j^r + BF_j^r(v, \delta d_2) \\ q_j^l &= [q_{11}, q_{12}]^T, F_j^l = \bar{M}_{11}\sigma(v_l) + \bar{M}_{12}\sigma(v_r) + \delta d_1 \\ q_j^r &= [q_{21}, q_{22}]^T, F_j^r = \bar{M}_{21}\sigma(v_l) + \bar{M}_{22}\sigma(v_r) + \delta d_2. \end{aligned} \quad (29)$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 06:18:28 UTC from IEEE Xplore. Restrictions apply.

Using the notations from (29), the tracking error dynamics of the wheels on left side can be expressed as:

$$\dot{e}_1 = A^o x_j + BF_j^l(v, q, x_d, r) - A_{dj}x_{dj} - B_{dj}r_j. \quad (30)$$

To exploit the Lemma 1 and quantify the tracking error bound, (30) is written as the following

$$\begin{aligned}\dot{e}_1 &= \mathbf{A^o x_j} + \mathbf{BF_j^l}(\mathbf{h_c}(.), \mathbf{q}, \mathbf{x_d}, \mathbf{r}) - \mathbf{A_{dj}x_{dj}} - \mathbf{B_{dj}r_j} \\ &- (A^o x_j + BF_j^l(h_c(.), q, x_d, r) - A_{dj}x_{dj} - B_{dj}r_j) \\ &+ A^o x_j + BF_j^l(v, q, x_d, r) + \delta d_1 - A_{dj}x_{dj} - B_{dj}r_j \quad (31)\end{aligned}$$

where the bold part (same as the reduced slow error dynamics (when $v = h_c(.)$)) represents the contracting part and the other terms can be considered as perturbation. Exploiting the smooth properties

$$\begin{aligned}&||F_j^l(v, q, x_d, r) - F_j^l(h_c(.), q, x_d, r) + \delta d_1|| \\ &\leq L_1||v(t) - h_c(.)|| + c_{lo}.\end{aligned} \quad (32)$$

where $L_1 \in \mathbb{R}^+$ is the associated Lipschitz constant in F corresponding to the left wheel dynamics. Similarly, let L_2 be the associated Lipschitz constant in F corresponding to the right wheel dynamics. Then, using the Lemma 1, the tracking error bounds (14) can be obtained.

REFERENCES

[1] A. A. Dobson, J. A. Marshall, and J. Larsson, "Admittance control for robotic loading: Design and experiments with a 1-tonne loader and a 14-tonne load-haul-dump machine," *Journal of Field Robotics*, vol. 34, no. 1, pp. 123–150, 2017.

[2] Y. Shi, M. R. Elara, A. V. Le, V. Prabakaran, and K. L. Wood, "Path tracking control of self-reconfigurable robot htetro with four differential drive units," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 3998–4005, 2020.

[3] "Johnston sweepers," https://www.johnstonsweepers.com/wp-content/uploads/2017/12/c201-e6-brochure-31506.pdf, accessed:2020/7/28.

[4] C. Wang and G. Parker, "Analysis of rotary brush control characteristics for a road sweeping robot vehicle," in *2014 International Conference on Mechatronics and Control (ICMC)*. IEEE, 2014, pp. 1799–1804.

[5] E. Prassler, D. Schwammkrug, B. Rohrmoser, and G. Schmidl, "A robotic road sweeper," in *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065)*, vol. 3. IEEE, 2000, pp. 2364–2369.

[6] T. T. Tun, L. Huang, R. E. Mohan, and S. G. H. Matthew, "Four-wheel steering and driving mechanism for a reconfigurable floor cleaning robot," *Automation in Construction*, vol. 106, p. 102796, 2019.

[7] M. M. Rayguru, S. Roy, and I. N. Kar, "Time-scale redesign-based saturated controller synthesis for a class of mimo nonlinear systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2019.

[8] G. Gungor, B. Fidan, and W. W. Melek, "Reconfigurable robot manipulators: adaptation, control, and mems applications," in *Advanced Mechatronics and MEMS Devices II*. Springer, 2017, pp. 169–194.

[9] A. A. Hayat, R. Parween, M. R. Elara, K. Parsuraman, and P. S. Kandasamy, "Panthera: Design of a reconfigurable pavement sweeping robot," in *2019 International Conf. on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 7346–7352.

[10] L. Yi, A. Vu Le, A. A. Hayat, C. S. C. S. Borusu, R. E. Mohan, N. H. K. Nhan, and P. Kandasamy, "Reconfiguration during locomotion by pavement sweeping robot with feedback control from vision system," *IEEE Access*, vol. 8, pp. 113 355–113 370, 2020.

[11] R. Parween, M. V. Heredia, M. M. Rayguru, R. E. Abdulkader, and M. R. Elara, "Autonomous self-reconfigurable floor cleaning robot," *IEEE Access*, vol. 8, pp. 114 433–114 442, 2020.

[12] A. Alamdari and V. Krovi, "Active reconfiguration for performance enhancement in articulated wheeled vehicles," in *Dynamic Systems and Control Conf.*, vol. 46193. American Society of Mechanical Engineers, 2014, p. V002T27A004.

[13] A. Alamdari, X. Zhou, and V. N. Krovi, "Kinematic modeling, analysis and control of highly reconfigurable articulated wheeled vehicles," in *International Design Engineering Technical Conf. and Computers and Information in Engineering Conf.*, vol. 55935. American Society of Mechanical Engineers, 2013, p. V06AT07A070.

[14] T. Jianguo, L. Xiong, Y. Fei, and D. Zongquan, "A wheel-arm reconfigurable mobile robot design and its reconfigurable configuration," in *2009 ASME/IFToMM International Conf. on Reconfigurable Mechanisms and Robots*. IEEE, 2009, pp. 550–557.

[15] M. Tarokh, "A unified kinematics modeling, optimization and control of universal robots: from serial and parallel manipulators to walking, rolling and hybrid robots," *Autonomous Robots*, pp. 1–16, 2020.

[16] D. Savitski, V. Ivanov, K. Augsburg, T. Emmei, H. Fuse, H. Fujimoto, and L. M. Fridman, "Wheel slip control for the electric vehicle with in-wheel motors: Variable structure and sliding mode methods," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 10, pp. 8535–8544, 2019.

[17] X. Ji, X. He, C. Lv, Y. Liu, and J. Wu, "Adaptive-neural-network-based robust lateral motion control for autonomous vehicle at driving limits," *Control Engineering Practice*, vol. 76, pp. 41–53, 2018.

[18] A. A. Hayat, R. Parween, M. R. Elara, K. Parsuraman, and P. S. Kandasamy, "Panthera: Design of a reconfigurable pavement sweeping robot," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 7346–7352.

[19] S. Roy, S. Nandy, R. Ray, and S. N. Shome, "Robust path tracking control of nonholonomic wheeled mobile robot: Experimental validation," *International Journal of Control, Automation and Systems*, vol. 13, no. 4, pp. 897–905, 2015.

[20] M. M. Rayguru, R. E. Mohan, R. Parween, L. Yi, A. V. Le, and S. Roy, "An output feedback based robust saturated controller design for pavement sweeping self-reconfigurable robot," *IEEE/ASME Transactions on Mechatronics*, 2021.

[21] C. Edwards and Y. B. Shtessel, "Adaptive continuous higher order sliding mode control," *Automatica*, vol. 65, pp. 183–190, 2016.

[22] S. Roy, J. Lee, and S. Baldi, "A new adaptive-robust design for time delay control under state-dependent stability condition," *IEEE Transactions on Control Systems Technology*, vol. 29, no. 1, pp. 420–427, 2020.

[23] S. Roy, S. Baldi, P. Li, and V. Narayanan, "Artificial-delay adaptive control for under-actuated euler-lagrange robotics," *IEEE/ASME Transactions on Mechatronics*, 2021.

[24] W. Lohmiller and J.-J. E. Slotine, "On contraction analysis for nonlinear systems," *Automatica*, vol. 34, no. 6, pp. 683–696, 1998.

[25] F. Forni and R. Sepulchre, "A differential lyapunov framework for contraction analysis," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 614–628, 2013.

[26] A. Pavlov, A. Pogromsky, N. van de Wouw, and H. Nijmeijer, "Convergent dynamics, a tribute to boris pavlovich demidovich," *Systems & Control Letters*, vol. 52, no. 3-4, pp. 257–261, 2004.

[27] D. Del Vecchio and J.-J. E. Slotine, "A contraction theory approach to singularly perturbed systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 3, pp. 752–757, 2012.

[28] M. M. Rayguru and I. N. Kar, "Contraction theory approach to disturbance observer based filtered backstepping design," *ASME. J. Dyn. Sys., Meas., Control.*, vol. 141, no. 8, pp. 084 501–084 508, 2019.

[29] J.-X. Xu, Z.-Q. Guo, and T. H. Lee, "Design and implementation of integral sliding-mode control on an underactuated two-wheeled mobile robot," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 7, pp. 3671–3681, 2013.

[30] M. M. Rayguru and I. N. Kar, "High gain observer-based saturated controller for feedback linearizable system," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 1, pp. 122–126, 2019.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 06:18:28 UTC from IEEE Xplore. Restrictions apply.

Soft Computing
https://doi.org/10.1007/s00500-021-06528-0

APPLICATION OF SOFT COMPUTING

Neuro-fuzzy network incorporating multiple lexicons for social sentiment analysis

Srishti Vashishtha[1] · Seba Susan[1]

Accepted: 2 November 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract
We have proposed MultiLexANFIS which is an adaptive neuro-fuzzy inference system (ANFIS) that incorporates inputs from multiple lexicons to perform sentiment analysis of social media posts. We classify tweets into two classes: neutral and non-neutral; the latter class includes both positive and negative polarity. This type of classification will be considered for applications that aim to test the neutrality of content posted by the users in social media platforms. In our proposed model, features are extracted by integrating natural language processing with fuzzy logic; hence, it is able to deal with the fuzziness of natural language in a very efficient and automatic manner. We have proposed a novel set of 64 rules for the proposed neuro-fuzzy network that can classify tweets correctly by working on fuzzy features fetched from VADER, AFINN and SentiWordNet lexicons. The proposed novel rules are domain independent, i.e., we can extend these rules for any textual data that employs lexicons. The antecedent and consequent parameters of the ANFIS are optimized by gradient descent and least squares estimate algorithms, respectively, in an iterative manner. The key contributions of this paper are: (1) a novel neuro-fuzzy system: MultiLexANFIS that takes as its input the positive and negative sentiment scores of tweets computed from multiple lexicons—VADER, AFINN and SentiWordNet, in order to classify the tweets into neutral and non-neutral content, (2) a novel set of 64 rules for the Sugeno-type fuzzy inference system—MultiLexANFIS, (3) single-lexicon-based ANFIS variants to classify tweets when multiple lexicons are not available and (4) comparison of MultiLexANFIS with different fuzzy, non-fuzzy and deep learning state of the art on various benchmark datasets revealing the superiority of our proposed neuro-fuzzy system for social sentiment analysis.

Keywords ANFIS · Tweets · Lexicon · Neuro-fuzzy network · Sentiment analysis · Social media

1 Introduction

Social sentiment analysis has a huge impact on daily life, in modern times, due to the widespread use of social media for expressing public opinion. People prefer to communicate on the online social network such as Twitter, Facebook, WhatsApp, Instagram and YouTube; among them, Twitter is the most popular microblogging forum for expressing public sentiment. Sentiment analysis (SA) is used to inspect the polarity or intention of tweets about any entity. The content of tweets could be movie reviews (Vashishtha and Susan 2021; Parveen and Pandey 2016; Rahman et al. 2016; Anagha et al. 2015; Kulkarni et al. 2014; Ghiassi et al. 2013), product reviews (Ghiassi et al. 2013; Indhuja and Reghu 2014), stock market trends and political events (Khan et al. 2020), political views (Katta and Hegde 2019), COVID-19 (Nemes and Kiss 2021) or about any other topic that attracts public opinion.

Tweets can be classified as positive, negative or neutral using either supervised or unsupervised machine learning algorithms. Generally, the twitter data contains a major portion of neutral tweets compared to positive and negative tweets. It has been proved that neutral samples are vital for learning polarity in order to obtain remarkable classification results (Koppel and Schler 2005, 2006). In this research work, classification of tweets is not the stereotype—positive, negative or neutral; rather, it is neutral and

✉ Srishti Vashishtha
srishtidtu@gmail.com

Seba Susan
seba_406@yahoo.in

[1] Delhi Technological University, Shahbad Daulatpur, Main Bawana Road, Delhi 110042, India

Published online: 29 November 2021

non-neutral, and the latter class includes both positive and negative polarity. The purpose of this type of classification will be beneficial for such systems that check neutral content posted by the users. There are many intelligence agencies that aim to track an individual's social media activities for national security reasons, in search of anti-nationalists or terrorist activities, etc. Social media is widely used by people, and they tend to post, like and share content of their interest on a regular basis. These agencies analyze the content posted by individuals to determine whether the content is neutral or biased toward any political event or leader, or generally indicative of any malicious intent. Hence, such tasks need to deploy SA on online social network in order to classify the social media posts into neutral and non-neutral category related to an individual or group. To the best of our knowledge, none of the previous works on SA have concentrated on solely classifying tweets into neutral and non-neutral and none of them have targeted to build models that would aid security applications based on sentiment scores. Also, mostly the authors use one lexicon at a time and have not simultaneously incorporated multiple lexicons into their learning framework as proposed in our work. Two types of ANFIS variants—single lexicon and multiple lexicons, are presented in this work. The architecture of the proposed MultiLexANFIS model consists of fuzzy-inspired feature extraction phase where the input features are fetched from multiple lexicons—VADER (Hutto and Gilbert 2014), AFINN (Nielsen 2011) and SentiWordNet (Baccianella et al. 2010), and fuzzy inference phase is constructed using a novel set of 64 rules devised specifically for these three lexicons. The rules can be easily extended for adaptation to a greater number of lexicons.

The major contributions of this research are: (1) development of a novel neuro-fuzzy system: MultiLexANFIS that integrates sentiment scores from multiple lexicons to classify tweets into neutral and non-neutral. These sentiment scores are computed from pre-processed textual data and are interpreted as a fuzzy membership values. (2) A novel set of 64 rules for Sugeno-type fuzzy inference system (FIS)—MultiLexANFIS, is constructed. Each lexicon provides two sentiment scores—positive and negative, so a total of six sentiment scores are derived from the VADER, AFINN and SentiWordNet lexicons for each tweet. Each of these six inputs is assigned two membership functions (MFs): *Low* and *High*. (3) We have also constructed single-lexicon-based ANFIS variants with nine rules each to classify tweets. Three variants that we call as VADER, AFINN and SentiWordNet-ANFIS consider the two sentiment scores of positive and negative as input at a time. Each of these inputs is assigned three membership functions (MFs): *Low, Medium* and *High*. These variants can be deployed when multiple lexicons are not available. The experiments are conducted on MultiLexANFIS and all three single-lexicon ANFIS variants; the results reveal that MultiLexANFIS performs better than all single-lexicon variants. (4) A comparative analysis of proposed MultiLexANFIS with fuzzy approaches and non-fuzzy approaches including deep learning models on various benchmark datasets reveals the supremacy of our proposed neuro-fuzzy system. The performance of all models and the state of the art is measured in terms of RMSE (root mean square error).

This paper is organized as follows. Section 2 describes several recent works related to ANFIS and sentiment analysis. The proposed MultiLexANFIS is explained in Sect. 3. The experimental setup and implementation details are discussed in Sect. 4. The results are demonstrated in Sect. 5. Finally, Sect. 6 draws the overall conclusion.

2 Related work

Sentiment analysis is a process that is involved in the determination and classification of opinions or feelings expressed in text, through the use of computing machines. The major challenge is to extract information from unstructured data. Most methodologies in this area focus on polarity or opinion classification into positive or negative classes (Siddiqua et al. 2016; Vashishtha and Susan 2020b; Jefferson et al. 2017; Vashishtha and Susan 2020a, b); some researchers also include a neutral class (Parveen and Pandey 2016; Anagha et al. 2015; Indhuja and Reghu 2014; Vashishtha and Susan 2018, 2019a). Others detect neutral opinions and filter out them to enhance binary polarity classification (Valdivia et al. 2017, 2018). Sentiment analysis models utilizing part-of-speech tags and n-gram feature extraction are gaining popularity. These models deploy either unsupervised algorithms (Vashishtha and Susan 2021) or supervised machine learning algorithms involving SVM, naïve Bayes, etc. (Tripathy et al. 2016). One notable supervised technique is the combination of hidden Markov model for feature extraction and SVM and ANN for the classification process (Kalarani and Brunda 2019). Machine learning classifier models have proved to perform well for sentiment classification, and some of the popular algorithms are naïve Bayes (Parveen and Pandey 2016), support vector machines (SVM) (Katta and Hegde 2019), (Valdivia et al. 2017, 2018), long short-term memory (LSTM) (Rahman et al. 2016; Vashishtha and Susan 2019), recurrent neural network (RNN) (Nemes and Kiss 2021), maximum entropy with probabilistic latent semantic analysis (Xie et al. 2019) and ANFIS. Out of all these options, neuro-fuzzy networks perform generally well in classification tasks involving ambiguous and

uncertain systems since they combine learning with reasoning which is a human trait (Susan et al. 2013).

2.1 Fuzzy logic-based sentiment analysis

In recent years, several fuzzy approaches have started to emerge for text processing. Fuzzy logic is a concept used in computing based on the "degree of truth" instead of true or false labels (0 or 1). It uses fuzzy sets, where each element belongs to the set with a certain degree; this degree is determined by a particular membership function. It has been meticulously exploited as a methodology useful both for representing information associated with sentiment and for executing various sentiment analysis tasks. The exploitation of fuzzy concepts in the field of SA can be incorporated in feature extraction, in fuzzy-based model architecture or in both (Serrano-Guerrero et al. 2021). An unsupervised fuzzy logic-based method for sentiment classification of textual reviews was proposed in Vashishtha and Susan (2020b). The authors formulated fuzzy cardinality, as the measure, from fuzzy sets for the evaluation of word polarity scores, and the proposed model has two versions based on the sentiment lexicon utilized in the model. Many research works have implemented fuzzy membership functions (MFs): triangular (Vashishtha and Susan 2019a), trapezoidal (Liu and Cocea 2017; Jefferson et al. 2017), Gaussian and generalized bell MFs (Gupta et al. 2014; Talpur et al. 2017). These MFs for a basic atomic term can be modified using linguistic Hegdes, which are adverbs or adjectives. Linguistic Hegdes with fuzzy logic have been successfully used to compute the sentiment of text (Vashishtha and Susan 2018, 2021; Khattak et al. 2020). A combination of fuzzy entropy, k-means clustering and sentiment lexicon shortlists words that help in sentiment cognition in Vashishtha and Susan (2019b). In the field of social media SA, there is also a shift toward multimodal social web, where users post their opinion in the form of text in Facebook and Twitter, audio or video clips in YouTube. The social media reviews are either classified using only fuzzy logic-based text classification (Wu et al. 2017) or can be classified by multimodal sentiment classification using a fusion of fuzzy logic with audio and text features (Vashishtha and Susan 2020a). Sentiment classification can be deployed using rule-based systems, a special type of expert system which uses a set of pre-defined rules for the classification task. Rule-based classifier can be supervised (Siddiqua et al. 2016) or may use fuzzy rules constructed in an unsupervised manner (Anagha et al. 2015; Liu and Cocea 2017; Jefferson et al. 2017; Vashishtha and Susan 2019a; Talpur et al. 2017). The rules can be further applied to supervised machine learning algorithms: naïve Bayes (Siddiqua et al. 2016) and SVM (Katta and Hegde 2019) and can also be learned in ANFIS (Jang 1993). Ahuja et al. (2015) implemented a SA model on public opinion data from Twitter to find a relationship between the stock market and public mood, using a self-organizing fuzzy neural network. An optimized fuzzy technique—shuffled frog leaping algorithm (SFLA), developed by Madhusudhanan et al. (2019) is deployed for improving the SA process. The SFLA is an optimization technique which imitates the mimetic evolutionary activity of frogs searching for a place with most quantity of food; along with SFLA, a local search algorithm is implemented for performing SA on book reviews. Phan et al. (2019) proposed a method that detects polarity of only specific kind of tweets that contain fuzzy sentiment phrases (FSPs). The first step of this method is to detect FSPs and then extract a set of features related to FSPs based on the syntactic, lexical, semantic and polarity sentiment of the words. The second step is to classify tweets using multilayer perceptron model. In a recent research work (Yakubu et al. 2021), customer satisfaction model based on online reviews has been developed, based on the sentiment scores derived by applying a novel multigene genetic programming-based fuzzy regression (MGGP-FR) approach. To address the fuzziness of customer opinions, the sentiment scores are transformed into asymmetrical fuzzy numbers. Then, the defuzzified sentiment scores are applied to MGGP; further, Tanaka's fuzzy regression is employed to determine whether the customer is satisfied or not.

2.2 Adaptive neuro-fuzzy inference system (ANFIS)

In 1993, Jang proposed ANFIS, a fuzzy inference system implemented in the framework of adaptive networks (Jang 1993). It uses a hybrid learning procedure to construct an input–output mapping based on both human knowledge (in the form of fuzzy if–then rules) and designated input–output data pairs. ANFIS has been applied for classification (Žunić et al. 2016), prediction (Cosma and Acampora 2016; Kulkarni et al. 2012, 2014; Mazumder et al. 2018), subjectivity analysis (Rustamov et al. 2013a; b; Rustamov 2018; Kamil et al. 2018; Padmaja and Hegde 2019) and SA (Katta and Hegde 2019), (Mazumder et al. 2018), (Kamil et al. 2018; Padmaja and Hegde 2019; Acampora and Cosma 2014; Çakıt et al. 2019). Kulkarni et al. (2012) developed an ANFIS based on sentiment information from movie blogs and reviews, for predicting sales performance in movie domain. The proposed predictive model is built to incorporate the auto-regressive method and ANFIS technique to handle the fuzzy demand with incomplete information. Another work by Kulkarni et al. (2012) proposed neuro-fuzzy methodology—ANFIS, for predicting sales performance of movies by analyzing sentiments in online reviews. The ANFIS model here has three inputs: review

ratings, revenue budget and type of the movies. Twitter data has been analyzed for political issues: using a hybrid system of ANFIS and SVM, where the SVM classifier optimizes the fuzzy rules (Katta and Hegde 2019), and using principal component analysis with ANFIS to predict political sentiments (Mazumder et al. 2018). Another hybrid classifier of ANFIS with genetic algorithm (GA) has shown good results in twitter SA (Padmaja and Hegde 2019; Acampora and Cosma 2014). GA extracts the important and semantic features from each tweet and optimizes the fuzzy principles in ANFIS. The authors in Asimuzzaman et al. (2017) implemented ANFIS to predict the polarity of Bangla tweets. The ANFIS utilizes semantic rules which are simple yet have a greatly affect the actual polarity of the text. Since the model is designed for classifying tweets into positive and negative polarity, the neutral tweets were ignored in pre-processing phase. Each and every word of Bangla text is translated into English text, so that the polarity value of each word can be extracted from SentiWordNet lexicon. The experiments of this model are conducted on a small twitter samples training set, and it has not been compared with recent state of the art for SA. Kamil et al. (2018) applied ANFIS for executing three different classification problems: (1) sentence-level subjectivity detection, (2) sentiment analysis of texts and (3) detecting user intention in natural language call routing system. In their work, the main goal is to avoid using human annotation or lexical knowledge; the membership degree of each term is calculated by pruned ICF (inverse-class frequency). The fuzzification operations applied involve the computation of the maximum of membership degree with respect to the classes for every term and mean of maxima for all classes. A recent study compared supervised soft computing approaches: fuzzy time series (FTS), artificial neural network (ANN)-based FTS and ANFIS to estimate the human emotional states expressed in tweets on social media networks (Çakıt et al. 2019). Combination of ANFIS and hidden Markov model has been deployed for SA in (Rustamov et al. 2013a; b) and subjectivity analysis in (Rustamov 2018). During the preparatory investigation in this study, it was found that triangular, Gaussian and generalized bell MFs have lower root mean square error (RMSE) and this motivated us to use these MFs in our proposed MultiLexANFIS. A multimodal convolutional neuro-fuzzy framework has been developed by Nguyen et al. (2019) that integrates the concepts of fuzzy logic with deep neural network for emotion classification of videos. Convolutional neural network is deployed to extract features from text, audio and visual modalities; these are sent as input to ANFIS that generates new rules to correctly classify emotions. An improved ANFIS (IANFIS) model has been proposed by Sasikala and Sheela (2020) aimed for future prediction of online products. The authors have also developed a deep learning modified neural network (DLMNN) for SA of online products. A food review dataset being analyzed, initially the dataset is divided into three groups: grade based, content based and collaboration based. For the grade-based data, polarity scores are determined and the review is classified as positive, negative or neutral using DLMNN, ANN and hybrid dragonfly-genetic algorithm (HDF-GA). Social media platform provides abundant and huge amount of data that can be analyzed for different purposes. A hybrid machine learning model: bi-objective optimization and LAN^{2}FIS (logistic adaptive network based on neuro-fuzzy inference system), proposed by Nagamanjula et al. (2020) aims to utilize and handle this vast amount of data for sentiment classification of tweets. The bi-objective optimization implements minimum redundancy and maximum relevancy for selecting optimal set of features from tweets. These features are fed as input to logistic regression-based ANFIS for initially classifying tweets into five classes: positive, negative, strongly positive, strongly negative and neutral; further, the classes are reduced to positive, negative and neutral.

3 Proposed ANFIS architecture for social sentiment analysis

In this paper, we have proposed a novel ANFIS-based system: MultiLexANFIS that integrates inputs from multiple sentiment lexicons to compute the sentiment of social media posts. We have also proposed single-lexicon-based ANFIS variants that require input from one sentiment lexicon at a time. These variants can be deployed when multiple lexicons are not available. The datasets chosen in this research work comprise tweets that are categorized into two classes: neutral and non-neutral. We have merged the positive and negative polarities into one class: non-neutral.

3.1 Motivation

In the last decade, fuzzy systems and their hybrid derivations have proved that they can emulate the typical human reasoning ability in a computationally efficient way. Fuzzy models have been used to construct cognitive models based on rules, expert systems, classifiers and universal approximators. Adaptive neuro-fuzzy inference system (ANFIS) is a neuro-fuzzy system that incorporates the power of artificial neural network and fuzzy logic. Our ANFIS model combines natural language processing (NLP) with fuzzy logic, which boosts the performance of SA. The importance of fuzziness comes into play while dealing with natural language due to the presence of ambiguity in word

senses and the semantic interpretation of sentences. The amalgamation of fuzzy logic concepts into SA helps in dealing with the fuzziness of natural language in a very efficient and automatic manner. Neuro-fuzzy systems have been applied, before, in the field of SA or opinion mining, but these systems were not able to handle the uncertainty and vagueness present in the natural language. Moreover, these systems usually classified the data into positive or negative class; our goal is to classify data into neutral on non-neutral class that would aid organizations to test for the neutrality of content posted by users in social media platforms about persons, events, services, etc. To the best of our knowledge, this novel sentiment classification problem—segregation of neutral and non-neutral classes, has not been explicitly handled by any research work. We have designed two types of ANFIS variants: for single lexicon and multiple lexicons. The single-lexicon variant is based on the type of sentiment lexicon chosen. The experiments have been designed to figure out which lexicon is best suited for ANFIS-based sentiment classification in the field of social SA. Since single-lexicon variant proved to be inconclusive about which fuzzy membership is best suited for ANFIS-based SA, we designed a multiple-lexicon ANFIS variant—MultiLexANFIS. To solve the above confusion of choosing the best membership function (MF) for a given lexicon and dataset combination, the MultiLexANFIS architecture has been proposed. This system integrates all the lexicons into a single unit and produces lower error compared to other models, and it takes less execution time in contrast to single-lexicon-based ANFIS. In our work, the proposed novel rules are domain independent. We can apply these rules for any textual data that employs lexicons to extract positive and negative score for each textual document or sentence or review.

3.1.1 Motivation behind segregating neutral and non-neutral classes

The motivation behind segregating neutral and non-neutral classes is that in some applications, the neutral content carries more importance than others. It has been demonstrated that a mix of training samples containing neutral samples is better as compared to a training set of the same size that includes only positives and negatives samples (Koppel and Schler 2005, 2006). The use of neutral training samples in learning not only promotes better distinction between positive and negative samples but also increases the classification accuracy (Koppel and Schler 2006). There exists a bias in social media content posted by users. Social theories claim that bias is a deep-seated feature of human behavior, which is described by a lack of proper balance and neutrality; one of the instances of bias is when someone supports one side too strongly or too often (Guerra et al. 2011). We have designed an ANFIS-based architecture that can be used by organizations which hunt for neutral and unbiased content. Currently, there is a new trend; tasks like personality profiling and behavior profiling are deployed by organizations based on the neutrality of content. Organizations which deal with visa applications may require the candidates to have neutral behavior for their country; those candidates who are too biased can pose a threat for that country. Such systems can also be used by military services and intelligence agencies for tracking anti-nationalists or terrorist activities.

3.1.2 Comparison of our proposed ANFIS model with existing ANFIS works in literature

Katta and Hegde (2019) proposed a hybrid ANFIS that integrates nonlinear SVM for SA of political twitter data. In feature extraction phase, unigrams and bigrams models are used. Only the words or pair of words are fed as input to ANFIS; whereas in our proposed model we evaluate the positive and negative sentiment scores of each tweet and send them as input to ANFIS. This is expected to yield better results because sentiment scores extracted from lexicons would provide better information for sentiment classification. In Katta and Hegde (2019), authors used only one fuzzy MF: generalized bell, while in our proposed model we have explored three fuzzy MFs: triangular (trimf), generalized bell (gbellmf) and Gaussian (gaussmf), in order to select the best performing MF. The authors in Katta and Hegde (2019) have conducted experiments and compared their model's different hybrid ANFIS-SVM variants but not with other state-of-the-art methods, whereas we have conducted experiments and compared with one fuzzy-rule-based classifier and four non-fuzzy state-of-the-art methods.

Kamil et al., (2018) applied ANFIS for executing three different classification problems. In the pre-processing phase, the text is converted into an array of words which is sorted in increasing order of alphabets. Each word is assigned a unique code. The words are not lemmatized; hence, neither verbs in different tenses nor nouns as singular or plural are combined. In our approach, the words are lemmatized. The advantage of lemmatization is that it takes into consideration the context of the word to determine which is the intended meaning the user is looking for. This process allows to decrease noise and speed up the user's task. The authors in Kamil et al. (2018) do not use lexical knowledge, whereas we use lexicon to extract the positive and negative polarity scores that contribute to SA. Sentiment lexicons permit the creation of important and high-quality features for supervised SA. Moreover, they can be applied to diverse domains. The incorporation of

sentiment lexicon makes our model domain independent. In Kamil et al. (2018), the membership degree of each term is calculated by pruned ICF (inverse-class frequency), whereas in our model, three popular fuzzy MFs are applied for fuzzification of sentiment scores.

Padmaja and Hegde (2019) developed a hybrid classifier based on ANFIS and genetic algorithm (GA) to classify the tweets into binary sentiment classes—positive or negative class, while our proposed model classifies tweets into neutral or non-neutral sentiment class. The ANFIS classifier integrates the benefits of both artificial neural network and fuzzy logic for generating the fuzzy rules, and the GA optimizes the fuzzy principles in the ANFIS classifier. After pre-processing of raw tweets, features are extracted; here, positive word count, negative word count and tag count are considered as features. The positive and negative words are attained by considering only adjectives, whereas in our proposed model, the positive and negative scores of tweets are computed by considering adjectives, nouns, verbs and adverbs. The inclusion of other part of speech enhances the sentiment evaluation process. Moreover, they have not specified how the words are classified as positive or negative. In our ANFIS model, we have deployed sentiment lexicons to compute the positive and negative scores of tweets. The authors in Padmaja and Hegde (2019) have compared with only one existing methodology latent Dirichlet allocation possibilistic fuzzy C-means (LDA-PFCM) and conducted experiments on only one dataset of 479 tweets. Our proposed model has been compared with eight existing SA approaches, and the experiments are conducted on ten benchmark datasets.

3.2 Processing of tweets

In this section, we will discuss the overall procedure of tweet processing. Initially, the twitter data is collected; many publicly available datasets consist of positive, negative and neutral tweets. We have used a total of ten publicly available datasets: Apple Twitter Dataset (2019), Nuclear Twitter Dataset (2019), Sanders Twitter Dataset (2019), SemEval 2017, SemEval 2016, SemEval 2015, Stanford Twitter Sentiment Test Set (STS-Test) (Go et al. 2009), Airline Twitter Dataset (2019), Twitter_2019 (2019) dataset and Reddit dataset. All the datasets contain three types of tweets: positive, negative and neutral. The class distribution of all datasets is represented in Table 3. We have combined the positive and negative tweets into one class: non-neutral; our goal is to classify the tweets by ANFIS into two classes: neutral and non-neutral. We change the labels of positive and negative tweets to non-neutral; thus, we have now two types of classes: neutral and non-neutral. The labels, 1: neutral and 0: non-neutral, are the class labels in ANFIS. The various steps in tweet processing are outlined in the following.

3.2.1 Data pre-processing

The first step is data pre-processing, which is very essential because raw tweets are in unstructured form. In data pre-processing, we eliminate URLs, "@" symbol and hashtag "#" symbol. We have re-phrased commonly used phrases (like "can't") with their grammatical form ("can not"). This has been re-phrased because words like "not" contain polarity scores, if these phrases are not replaced, these words would be ignored which can lead to incorrect results.

3.2.2 Feature extraction using sentiment lexicons

We have to compute the sentiment of this pre-processed tweet which is required to be in a numerical format so that it can be used as an input to ANFIS. Sentiment lexicons are employed for this purpose, and these lexicons assign scores to either words or the whole tweet. The three sentiment lexicons are VADER (Hutto and Gilbert 2014), AFINN (Nielsen 2011) and SentiWordNet (Baccianella et al. 2010). These sentiment lexicons are popularly used in SA applications.

VADER (Valence Aware Dictionary for sEntiment Reasoning) is a sentiment lexicon that is especially attuned to sentiment in microblog-like contexts (Hutto and Gilbert 2014). It works fast and yields high accuracy. It computes the scores for the whole tweet s using VADER lexicon's *polarity_scores(s)* method and gives the positive (*TweetPos*) and negative (*TweetNeg*) scores of a tweet as the output. AFINN lexicon consists a list of English terms manually rated for valence with an integer between -5 (negative) and $+5$ (positive) (Nielsen 2011). This list also contains slang and obscene words; hence, it can deal with modern-day tweets very well. AFINN lexicon employs the AFINN method to fetch the score of each word a as represented in Eq. (1). If the score is greater than 0, it is a positive word (Eq. 2). If the score is less than 0, it is a negative word (Eq. 3). The positive score of the tweet (TweetPos) is calculated by adding up the scores of all the positive words in the tweet as shown in Eq. (4). The negative score of the tweet (TweetNeg) is calculated by adding up the scores of all the negative words in the tweet as shown in Eq. (5).

$$\text{afn} = \text{afinn.score}(a) \tag{1}$$

$$\text{if}(\text{afn} > 0)\text{then}(\text{af}_{\text{Pos}}(a) = \text{afn})) \tag{2}$$

$$\text{if}(\text{afn} < 0)\text{then}(\text{afNeg}(a) = -\text{afn}) \tag{3}$$

$$\text{TweetPos} = \sum_{a=1}^{m} \text{af}_{\text{Pos}}(a) \tag{4}$$

$$\text{TweetNeg} = \sum_{a=1}^{m} \text{af}_{\text{Neg}}(a) \tag{5}$$

where a is a word in a tweet and m is the number of selected words.

SentiWordNet lexicon is one of the most popular sentiment lexicons used by researchers. It assigns three numerical scores to each word: positive, negative and objective; these scores are in the range of 0.0 to 1.0 (Nielsen 2011). The tweets are processed for this lexicon in the following manner: We remove the stopwords and punctuations, perform lemmatization, part-of-speech (POS) tagging and word sense disambiguation (WSD) by Lesk (Banerjee and Pedersen 2002). WSD is the method of determining which "sense" (meaning) of a word is activated by the use of the word in a particular context. The SentiWordNet method acquires the scores of each word a from this lexicon using syn.pos_score () and syn.neg_score (). Each word has positive (Eq. (8)) and negative score (Eq. (9)) computed using WSD, which can be interpreted as a fuzzy membership belonging to the fuzzy sets Pos (Eq. (6)) and Neg (Eq. (7)). The words, which are having higher positive score than negative score in a tweet, are summed up to evaluate the positive score (TweetPos) of the tweet (Eq. (10)). Similarly, words which are having higher negative score than positive score in a tweet are summed up to evaluate the negative score (TweetNeg) of the tweet (Eq. (11)). These scores are computed for all tweets.

$$\text{Pos} = \{(a, \mu_{\text{Pos}}(a)\}, a \in Xi \tag{6}$$

$$\text{Neg} = \{(a, \mu_{\text{Neg}}(a)\}, a \in Xi \tag{7}$$

$$\mu_{\text{Pos}}(a) = \text{syn.pos\_score}() \tag{8}$$

$$\mu_{\text{Neg}}(a) = \text{syn.neg\_score}() \tag{9}$$

$$\text{if}(\mu_{\text{Pos}}(a) > 0 \& \mu_{\text{Pos}}(a) > \mu_{\text{Neg}}(a)) \quad \text{then}\left(\text{TweetPos} = \sum_{a=1}^{m} \mu_{\text{Pos}}(a)\right) \tag{10}$$

$$\text{if}(\mu_{\text{Neg}}(a) > 0 \& \mu_{\text{Neg}}(a) > \mu\text{Pos}(a)) \quad \text{then}\left(\text{TweetNeg} = \sum_{a=1}^{m} \mu_{\text{Neg}}(a)\right) \tag{11}$$

where a is the word in a tweet, m is the number of selected words and X_i is the set of total words. The computed TweetPos and TweetNeg of each tweet represent the numerical score of that tweet; these scores are given as input in ANFIS.

3.3 ANFIS processing of tweets

ANFIS was introduced by Jang in (1993); it takes crisp inputs and fuzzifies these inputs using membership functions. ANFIS is a method of artificial intelligence that uses a hybrid way of learning by connecting fuzzy logic and neural networks into a single unit. It is based on specific "human features," because fuzzy logic has a human subjectivity feature, while the neural network has a feature of human way of learning from past experience. ANFIS network is appropriate for solving function problems in various fields. Generally, the problem to be solved is the approximation of an unknown function. We next explain the working of a general ANFIS network constituted by M rules of Sugeno first-order type (Sugeno and Kang 1988). The kth rule, $k = 1 \ldots M$, has the following form:

$$\text{if } x_1 \text{ is } B_1^{(k)}, \text{and } x_2 \text{ is } B_2^{(k)} \ldots x_n \text{ is } B_n^{(k)}, \text{then } y^{(k)} = \sum_{j=1}^{n} a_j^{(k)} x_j + a_0^{(k)} \tag{12}$$

where $x = [x_1 x_2 \ldots x_n]$ is the input pattern and $y^{(k)}$ is the output associated with the kth rule. Each rule is characterized by the membership functions (MFs) $\mu_{B_j^{(k)}}(x_j)$ of the fuzzy input variables $B_j^{(k)}$, j = 1…n, and by the coefficients $a_j^{(k)}$, j = 1…n, of the crisp output. There are numerous options available for the fuzzification of crisp inputs, the composition of input MFs, and the way rule outputs are combined. The next step is to find the AND product of the fuzzy memberships of all input variables in a rule. This product is a MF $\mu_{B^{(k)}}(x)$, also called the firing strength (or weight) of the rule. The M-rule fuzzy inference system for ANFIS networks can be mathematically represented as:

$$\tilde{y} = \frac{\sum_{k=1}^{M} \mu_{B^{(k)}}(x) \cdot y^{(k)}}{\sum_{k=1}^{M} \mu_{B^{(k)}}(x)} \tag{13}$$

where $\tilde{y}$ is the estimate of the actual value $y = f(x)$, for a given input x.

We have proposed two ANFIS-based architectures: One is the single-lexicon-based ANFIS and the other, MultiLexANFIS that combines all the lexicons into a single integrated unit. Single-lexicon-based ANFIS takes two inputs, TweetPos and TweetNeg, and one output, sentiment label (1: neutral or 0: non-neutral), for one sentiment lexicon–dataset combination at a time. MultiLexANFIS has six inputs; it includes the positive and negative score of a tweet from all the three lexicons: VADER, AFINN and SentiWordNet, and one sentiment label as output for each dataset.

3.3.1 Single-lexicon-based ANFIS

The fuzzy inference system (FIS) is of Sugeno first-order type with two inputs TweetPos and TweetNeg and one output label *sentiment*. A common rule set with two inputs and three MFs for each input would yield nine rules ($3^2 = 9$). The MFs for the input have linguistic labels: *Low*, *Medium* and *High*. The labels *Low*, *Medium* and *High* are two fuzzy sets created using different fuzzy membership for input variables: positive score (TweetPos*)* and negative score (TweetNeg*)*. Mathematically, we can denote the inputs as x and y, MFs for input x as A_i, MFs for input y as B_i, where $i = 1, 2, 3$. The fuzzy if–then rules are then defined as follows:

$$\text{Rule1}: \text{If } x \text{ is} A_1 \text{ and } y \text{ is } B_1 \text{ then } f_1 = p_1x + q_1y + r_1 \tag{14}$$

$$\text{Rule2}: \text{If } x \text{ is } A_2 \text{ and } y \text{ is } B_1 \text{ then } f_2 = p_2x + q_2y + r_2 \tag{15}$$

$$\text{Rule3}: \text{If } x \text{ is } A_3 \text{ and } y \text{ is } B_1 \text{ then } f_3 = p_3x + q_3y + r_3 \tag{16}$$

$$\text{Rule4}: \text{If } x \text{ is } A_1 \text{ and } y \text{ is } B_2 \text{ then } f_4 = p_4x + q_4y + r_4 \tag{17}$$

$$\text{Rule5}: \text{If } x \text{ is } A_2 \text{ and } y \text{ is } B_2 \text{ then } f_5 = p_5x + q_5y + r_5 \tag{18}$$

$$\text{Rule6}: \text{If } x \text{ is } A_3 \text{ and } y \text{ is } B_2 \text{ then } f_6 = p_6x + q_6y + r_6 \tag{19}$$

$$\text{Rule7}: \text{If } x \text{ is } A_1 \text{ and } y \text{ is } B_3 \text{ then } f_7 = p_7x + q_7y + r_7 \tag{20}$$

$$\text{Rule8}: \text{If } x \text{ is } A_2 \text{ and } y \text{ is } B_3 \text{ then } f_8 = p_8x + q_8y + r_8 \tag{21}$$

$$\text{Rule9}: \text{If } x \text{ is } A_3 \text{ and } y \text{ is } B_3 \text{ then } f_9 = p_9x + q_9y + r_9. \tag{22}$$

The antecedent part of these rules is the same as in the recently published previous work of authors on unsupervised fuzzy rule-based sentiment analysis system (Vashishtha and Susan 2019a). In Mamdani FIS, a fuzzy system with two non-interactive inputs A and B (antecedents) and a single output C (consequent) is expressed by a number of r linguistic IF–THEN propositions in the Mamdani form:

$$\text{Rule } R_j: \text{IF } A \text{ is } A_1^j \text{ and } B \text{ is } B_1^j \text{ THEN } C \text{ is } C_1^j, \quad j = 1, 2, \ldots, r \tag{23}$$

where A_{1j} and B_{1j} are the fuzzy sets representing the jth antecedent or premise pairs and C_{1j} is the fuzzy set representing the jth consequent.

In this work for single-lexicon ANFIS, we transform the set of nine Mamdani fuzzy rules in (Vashishtha and Susan 2019a) to parametrized Sugeno form where the parameters are learnt by ANFIS. Mamdani FIS entails a substantial computational burden and is well suited to human input, while Sugeno FIS is computationally efficient, has flexibility in system design and is well appropriate to mathematical analysis. Hence, our proposed system in this work applies Sugeno FIS for better classification results.

The ANFIS architecture comprises five layers (Jang et al. 1997). Our single-lexicon ANFIS architecture is shown in Fig. 1. We denote the output of the ith node in layer l as $O_{l,i}$. Each layer is described next.

Layer 1 Every node i in this layer is an adaptive node with a node function

$$O_{l,i} = \mu_{A_i}(x), \quad \text{for } i = 1, 2, 3 \text{ or} \tag{24}$$

$$O_{l,i} = \mu_{B_{i-3}}(y), \quad \text{for } i = 4, 5, 6 \tag{25}$$

where $O_{l,i}$ is the membership grade of a fuzzy set A (= A_1, A_2 or A_3) or B (= B_1, B_2 or B_3). Here, the MF for A and B can be any parametrized MF: triangular (trimf), generalized bell (gbellmf) and Gaussian (gaussmf) described in Eqs. (26), (27) and (28) respectively.

$$\mu_t(x) = \begin{cases} 0, & x \le d \\ (x-d)/(e-d), & d < x \le e \\ (f-x)/(f-e), & e < x \le f \\ 0, & x \ge f \end{cases} \tag{26}$$

$$\mu_{gb}(x) = \frac{1}{1 + \left|\frac{x-w}{u}\right|^{2v}} \tag{27}$$

$$\mu_{ga}(x) = e^{-\frac{1}{2}\left(\frac{x-c}{\sigma}\right)^2}. \tag{28}$$

A brief explanation of these MFs ensues. The triangular membership function has the simplest shape among others. It is defined by three parameters: d and f for feet and e for the tip of the curve. The mathematical formula of the function is given in Eq. (26). Gbellmf is a symmetrical shape similar to a bell. As expressed by Eq. (27), this function employs three parameters: u determines the width of the bell like curve, v is a positive integer, while w sets the center of the curve in the universe of discourse. Gaussmf also has a smooth curve. However, as compared to all MFs mentioned above, it utilizes only two parameters: c for locating center and σ for determining the width of the curve as expressed mathematically by Eq. (28). The values of these parameters are attuned in the learning phase of ANFIS. The parameters in this layer are referred as premise parameters.

Layer 2 Every node in this layer is a fixed node labeled as $\prod$, whose output is the product of all the incoming signals shown in Eq. (29). Each node output represents the firing strength of a rule, where i = 1, 2,…, 9.

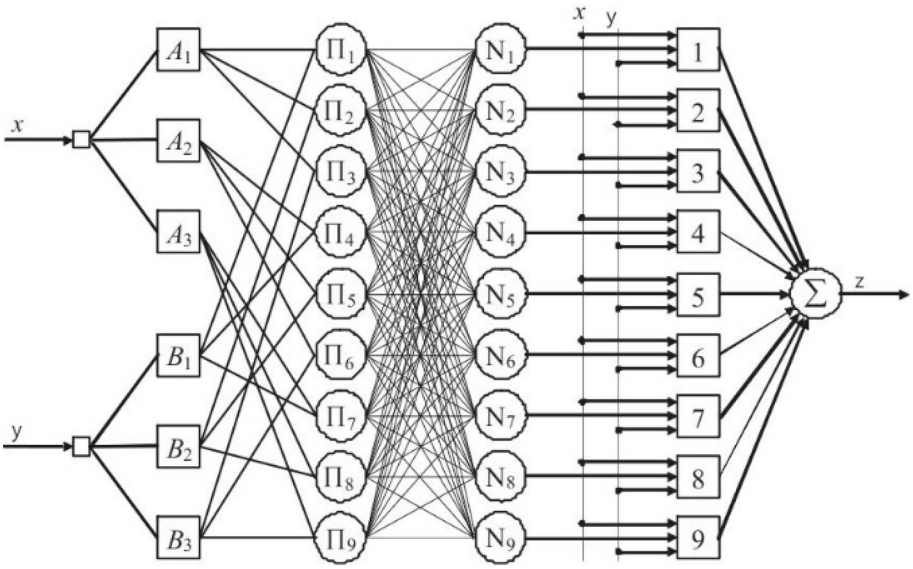


Fig. 1 Single-lexicon-based ANFIS architecture with nine rules

$$O_{2,i} = w_i = \mu_{A_k}(x)\mu_{B_j}(y), \quad k, j = 1, 2, 3. \tag{29}$$

Layer 3 Every node in this layer is a fixed node labeled *N*. The *i*th node calculates the ratio of the *i*th rule's firing strength to the sum of all rules' firing strengths in Eq. (30). Outputs of this layer are called the normalized firing strengths.

$$O_{3,i} = \overline{w} = \frac{w_i}{\sum_i w_i}. \tag{30}$$

Layer 4 Every node i in this layer is an adaptive node with a node function:

$$O_{4,i} = \overline{w_i} f_i = \overline{w}_i(p_i x + q_i y + r_i) \tag{31}$$

where w_i is the normalized firing strength from layer 3 and $\{p_i, q_i, r_i\}$ is the parameter set of this node. Parameters in this layer are referred to as consequent parameters.

Layer 5 The single node in this layer is a fixed node labeled $\sum$, which computes the overall output as the summation of all incoming signals:

$$\text{overall output} = O_{5,i} = \sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}. \tag{32}$$

3.3.2 Multiple-lexicon-based ANFIS: MultiLexANFIS

MultiLexANFIS combines all the three lexicons into a single integrated unit. MultiLexANFIS has six inputs; it includes positive and negative scores of a tweet from all the three lexicons, and one sentiment label as output for each dataset. The six inputs are: *vpos*, *vneg*, *apos*, *aneg*, *spos* and *sneg*. *vpos* and *vneg* are positive and negative score of a tweet computed using VADER lexicon (Hutto and Gilbert 2014); similarly, *apos* and *aneg* are evaluated using AFINN lexicon (Nielsen 2011) and *spos* and *sneg* are calculated using SentiWordNet lexicon (Baccianella et al. 2010) for the same tweet. The output is the sentiment label, which has value 1 for neutral tweet and 0 for non-neutral tweet. For ease of computations and for reducing the number of fuzzy rules, each input has been assigned two membership functions (MFs): *Low* and *High*. The labels *Low* and *High* are two fuzzy sets created using different fuzzy membership for input variables: *vpos*, *vneg*, *apos*, *aneg*, *spos* and *sneg*. Hence, there are a total of 64 rules (2^6 = 64). These rules are depicted in Tables 1 and 2. These proposed novel rules are domain independent. We can apply these rules for any textual data that employs lexicons to extract positive and negative scores for each textual document or sentence or review. MultiLexANFIS architecture has five layers similar to the single-lexicon ANFIS discussed in Sect. 3.3.1, except for the fact that the number of inputs has increased from two to six and the number of MFs for each input is two. MultiLexANFIS architecture is depicted in Fig. 2.

ANFIS learns by updating tunable parameters which are the premise parameters and the consequent parameters. The first and fourth layers have adaptive nodes, so they are trainable, whereas the nodes of rest of the layers are fixed. The optimization method adopted by ANFIS is, hybrid learning algorithm, composed of two passes.

The first pass is known as forward pass; here, ANFIS computes node outputs until the fourth layer where it uses least square estimate method to update the consequent parameters before calculating the final output. In the second pass or backward pass, the error is propagated

Table 1 MultiLexANFIS rules (1–32)

Rules	Input1	Input2	Input3	Input4	Input5	Input6	Output
1	Low	Low	Low	Low	Low	Low	O1
2	Low	Low	Low	Low	Low	High	O2
3	Low	Low	Low	Low	High	Low	O3
4	Low	Low	Low	Low	High	High	O4
5	Low	Low	Low	High	Low	Low	O5
6	Low	Low	Low	High	Low	High	O6
7	Low	Low	Low	High	High	Low	O7
8	Low	Low	Low	High	High	High	O8
9	Low	Low	High	Low	Low	Low	O9
10	Low	Low	High	Low	Low	High	O10
11	Low	Low	High	Low	High	Low	O11
12	Low	Low	High	Low	High	High	O12
13	Low	Low	High	High	Low	Low	O13
14	Low	Low	High	High	Low	High	O14
15	Low	Low	High	High	High	Low	O15
16	Low	Low	High	High	High	High	O16
17	Low	High	Low	Low	Low	Low	O17
18	Low	High	Low	Low	Low	High	O18
19	Low	High	Low	Low	High	Low	O19
20	Low	High	Low	Low	High	High	O20
21	Low	High	Low	High	Low	Low	O21
22	Low	High	Low	High	Low	High	O22
23	Low	High	Low	High	High	Low	O23
24	Low	High	Low	High	High	High	O24
25	Low	High	High	Low	Low	Low	O25
26	Low	High	High	Low	Low	High	O26
27	Low	High	High	Low	High	Low	O27
28	Low	High	High	Low	High	High	O28
29	Low	High	High	High	Low	Low	O29
30	Low	High	High	High	Low	High	O30
31	Low	High	High	High	High	Low	O31
32	Low	High	High	High	High	High	O32

Table 2 MultiLexANFIS rules (33–64)

Rules	Input1	Input2	Input3	Input4	Input5	Input6	Output
33	High	Low	Low	Low	Low	Low	O33
34	High	Low	Low	Low	Low	High	O34
35	High	Low	Low	Low	High	Low	O35
36	High	Low	Low	Low	High	High	O36
37	High	Low	Low	High	Low	Low	O37
38	High	Low	Low	High	Low	High	O38
39	High	Low	Low	High	High	Low	O39
40	High	Low	Low	High	High	High	O40
41	High	Low	High	Low	Low	Low	O41
42	High	Low	High	Low	Low	High	O42
43	High	Low	High	Low	High	Low	O43
44	High	Low	High	Low	High	High	O44
45	High	Low	High	High	Low	Low	O45
46	High	Low	High	High	Low	High	O46
47	High	Low	High	High	High	Low	O47
48	High	Low	High	High	High	High	O48
49	High	High	Low	Low	Low	Low	O49
50	High	High	Low	Low	Low	High	O50
51	High	High	Low	Low	High	Low	O51
52	High	High	Low	Low	High	High	O52
53	High	High	Low	High	Low	Low	O53
54	High	High	Low	High	Low	High	O54
55	High	High	Low	High	High	Low	O55
56	High	High	Low	High	High	High	O56
57	High	High	High	Low	Low	Low	O57
58	High	High	High	Low	Low	High	O58
59	High	High	High	Low	High	Low	O59
60	High	High	High	Low	High	High	O60
61	High	High	High	High	Low	Low	O61
62	High	High	High	High	Low	High	O62
63	High	High	High	High	High	Low	O63
64	High	High	High	High	High	High	O64

backward until the first layer where ANFIS employs gradient descent to tune the premise parameters. In forward pass, the antecedent part is fixed, while in backward pass the consequent part is fixed. The cost of training and computational complexity of the network is ascertained by the apt choice of the number and shape of membership functions.

The performance of ANFIS is measured using root mean square error (RMSE). It is the most popularly used measure. It is a measure of the differences between values predicted by the model and the observed values, as shown in Eq. (33). In other words, it measures the quality of the fit between the actual data and the predicted model.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{t=1}^{n} (\mathrm{actual}(t) - \mathrm{predicted}(t))^2}{p}} \tag{33}$$

where actual*(t)* and predicted*(t)* are the actual and estimated values, respectively, and n is the total number of records in the testing data, $t = 1,2,3\ldots p$. The overall process flow of our proposed ANFIS for social SA is summarized in Fig. 3.

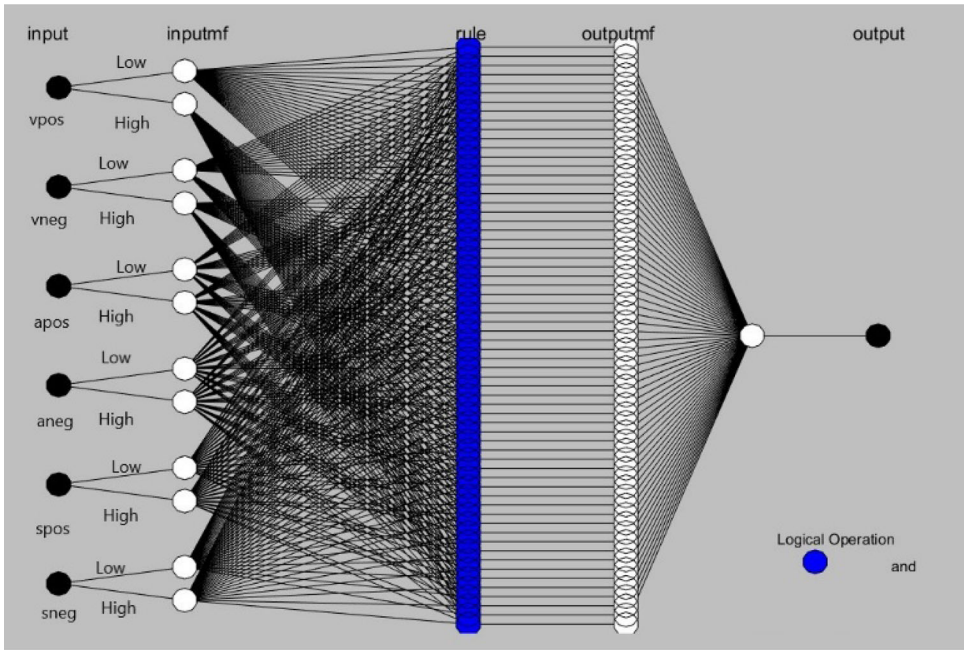


Fig. 2 MultiLexANFIS architecture with 64 rules

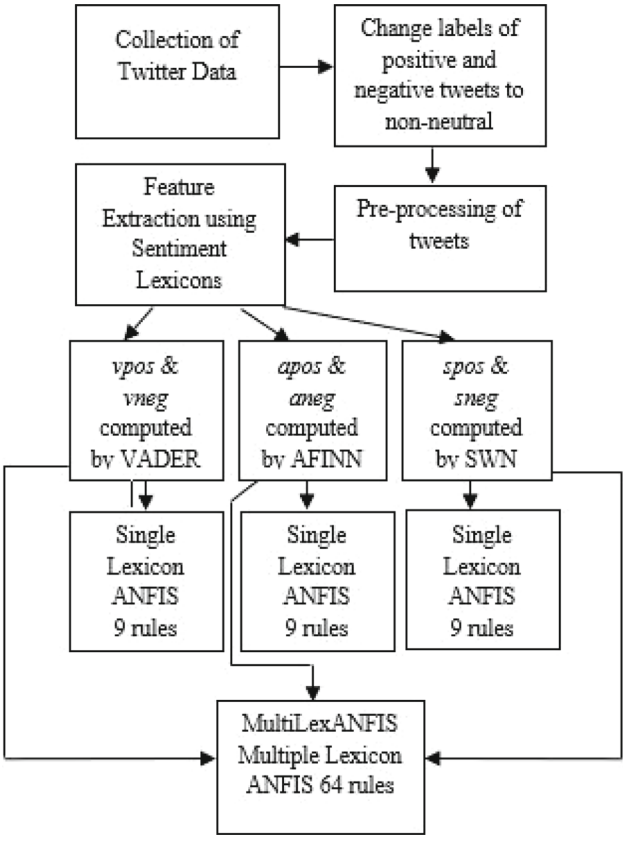


Fig. 3 Overall process of our proposed ANFIS for social SA

4 Experimental setup and implementation

In the experimental phase, the proposed ANFIS was simulated in MATLAB 2020b version with Intel Core i5 processor, 64-bit operating system and 8 GB RAM. We have used a total of ten publicly available datasets: Apple Twitter Dataset (2019), Nuclear Twitter Dataset (2019), Sanders Twitter Dataset (2019), SemEval 2017, SemEval 2016, SemEval 2015, Stanford Twitter Sentiment Test Set (STS-Test) (Go et al. 2009), Airline Twitter Dataset (2019), Twitter_2019 (2019) dataset and Reddit dataset. All the datasets contain three types of tweets: positive, negative and neutral. The class distribution of all datasets is represented in Table 3. We have combined the positive and negative tweets into one class: non-neutral; our goal is to classify the tweets by ANFIS into two classes: neutral and non-neutral. The training and testing sets of these datasets are constructed using 70–30% split ratio. Our proposed model uses inputs derived from sentiment lexicons. The three sentiment lexicons used in this study are: VADER (Hutto and Gilbert 2014), AFINN (Nielsen 2011) and SentiWordNet (Baccianella et al. 2010). Our proposed model is compared with eight state-of-the-art methods that include both fuzzy and non-fuzzy approaches; their description is mentioned in the next section.

In our single-lexicon-based ANFIS variants, there are two inputs: positive and negative scores, and one output for sentiment. Class 1 indicates neutral tweets, and Class 0 indicates non-neutral tweets. The number of fuzzy rules in grid partitioning is number of MFs ^ number of input variables. In single-lexicon-based ANFIS variant, since

Table 3 Class distribution of different datasets

	Sanders	Nuclear	Apple	SemEval 2017	SemEval 2016	SemEval 2015	STS-Test (Go et al. 2009)	Airline	Twitter	Reddit
Positive	519	10	423	2375	5157	4377	182	2363	72,249	15,829
Negative	572	19	1219	3972	1225	1745	177	9178	35,509	8278
Neutral	2333	161	2162	5937	2667	5593	139	3099	55,212	12,909
Total	3424	190	3804	12,284	9049	11,715	498	14,640	162,970	37,014

each input has three membership functions (MFs), so in total, ANFIS has nine rules (3^2 = 9). In MultiLexANFIS, there are six inputs (two from each of the three lexicons) and the same output variable for sentiment, and each input has two MFs. Hence, it has a total of 64 rules (2^6 = 64). Optimal number of inputs and MFs are selected for both the cases. In MultiLexANFIS, we have considered two MFs compared to three MFs in single-lexicon ANFIS variants in order to reduce the curse of dimensionality. The "curse of dimensionality" refers to such situation where the number of fuzzy rules, when the grid partitioning is used, increases exponentially. Grid partitioning generates rules by enumerating all possible combinations of membership functions of all inputs; this leads to an exponential explosion even when the number of inputs is moderately large. ANFIS with grid partitioning was used to test different MFs: triangular, Gaussian and generalized bell. The optimization method used to train the ANFIS is hybrid learning, which combines the gradient descent method and the least squares estimate (LSE) to learn the parameters. For all the variants, the error tolerance is set to 0. While conducting experiments, it was observed that training error in single-lexicon ANFIS variants remains constant after 500 epochs, while in MultiLexANFIS, the training error remains constant after 100 epochs. Hence, the number of epochs required by single lexicon is 500 epochs, while MultiLexANFIS requires 100 epochs during training phase.

5 Results and discussion

In this paper, we have proposed two ANFIS for sentiment analysis of social media data: single-lexicon-based ANFIS and MultiLexANFIS, that incorporates inputs from multiple lexicons. These systems have been implemented on ten datasets. Root mean square error (RMSE) is the mostly used metric in most of the papers involving ANFIS models. Thus, we have used this as the performance metric in all experiments. All the RMSE values depicted in this work are test errors.

Our ANFIS has been compared with eight state-of-the-art methods that include various fuzzy, non-fuzzy and deep learning models. The eight state-of-the-art methods are: (1) A deep learning model—recurrent neural network (RNN), developed by (Nemes and Kiss 2021) for classifying tweets. (2) A deep learning model—long short-term memory (LSTM) model, developed by (Rahman et al. 2016) for text classification. (3) An n-gram model developed by (Tripathy et al. 2016) that converts text reviews into numeric matrices using count vectorizer and TF-IDF, which are then given as input to machine learning algorithm—support vector machine (SVM) for classification. (4) An unsupervised fuzzy rule-based system for sentiment analysis is a published previous work of authors (Vashishtha and Susan 2019a). This fuzzy rule-based classifier has nine rules implemented in a Mamdani fuzzy inference system. (5) VADER SA by Hutto and Gilbert (2014); this approach yields positive, negative, neutral scores for a given text by deploying a number of rules. (6) AFINN SA by Nielsen (2011) approach uses AFINN dictionary to assign scores to each word; the sentiment is computed for text by summing up all scores. (7) SA using SentiWordNet lexicon by Cavalcanti et al. (2011) approach uses SentiWordNet to assign scores to selected words—adjectives, verbs, adverbs and nouns, and then, final sentiment is computed by considering the average score of all words. (8) SA using SentiWordNet lexicon by Ortega et al. (2013) approach uses SentiWordNet to check the polarity of words, assigns fixed values for each polarity and then computes the sentiment by summing up the values.

5.1 Single-lexicon-based ANFIS

In this research work, we have considered three sentiment lexicons, so there are three single-lexicon-based ANFIS: VADER (Hutto and Gilbert 2014), AFINN (Nielsen 2011) and SentiWordNet (Baccianella et al. 2010) single-lexicon-based ANFISs. These systems have nine rules. We have performed experiments with each of these ANFISs for all the datasets and three MFs; the aim is to find out which MF is best suitable for the given ANFIS. Table 4 depicts the outcomes for VADER lexicon-based ANFIS, and the bold values highlight the lowest RMSE achieved for each dataset. Both trimf and gaussmf membership functions

Table 4 VADER lexicon-based ANFIS

Datasets	TRIMF	GBELL	GAUSS
	Root mean square error (RMSE)		
Apple	0.46213	**0.45668**	0.45887
Nuclear	0.37404	0.35645	**0.33174**
Sanders	**0.43453**	0.43458	0.43683
SemEval (2017)	0.47819	0.47825	**0.47814**
SemEval (2016)	**0.45195**	0.45249	0.45197
SemEval (2015)	**0.45257**	0.45301	0.45266
STS (Go et al. 2009)	0.37445	0.36976	**0.36864**
Airline	0.39421	0.38901	**0.388**
Twitter	0.44973	**0.43362**	0.43371
Reddit	0.43592	**0.39184**	0.39223

Bold values are the lowest RMSE values for each dataset

Table 5 AFINN lexicon-based ANFIS

Datasets	TRIMF	GBELL	GAUSS
	Root mean square error (RMSE)		
Apple	0.48046	0.50781	**0.46623**
Nuclear	0.4105	**0.39682**	0.40527
Sanders	**0.44414**	0.45402	0.44803
SemEval (2017)	**0.47428**	0.47486	0.47449
SemEval (2016)	0.45577	**0.45331**	0.45535
SemEval (2015)	0.44567	0.44541	**0.44535**
STS (Go et al. 2009)	0.40142	**0.38946**	0.39323
Airline	0.38994	**0.38732**	0.38982
Twitter	0.42823	**0.42691**	0.42991
Reddit	0.43172	**0.42713**	0.45053

Bold values are the lowest RMSE values for each dataset

Table 6 SENTIWORDNET lexicon-based ANFIS

Datasets	TRIMF	GBELL	GAUSS
	Root mean square error (RMSE)		
Apple	0.4878	**0.48503**	0.4866
Nuclear	**0.36351**	0.36638	0.36921
Sanders	0.45498	0.45341	**0.45318**
SemEval (2017)	0.49336	**0.49243**	0.49294
SemEval (2016)	**0.45204**	0.4526	0.45576
SemEval (2015)	0.48483	**0.4844**	0.48484
STS (Go et al. 2009)	**0.42304**	0.42831	0.42352
Airline	0.39752	**0.39614**	0.39623
Twitter	0.43542	0.43032	**0.43214**
Reddit	**0.45023**	0.46493	0.45302

Bold values are the lowest RMSE values for each dataset

(MFs) have lower RMSE in a greater number of datasets compared to gbellmf. The least RMSE, 0.33174, is scored in nuclear dataset under gaussmf, while in the case of AFINN-based ANFIS in Table 5, gbellmf has performed better than trimf and gaussmf; it has the lowest RMSE for three datasets. The least RMSE, 0.3873, is scored in Airline Twitter dataset under gbellmf. In Table 6, we can observe the results for SentiWordNet-based ANFIS; here, trimf and gbellmf have more cases of lower RMSE compared to gaussmf. Again, the Nuclear dataset has least RMSE of 0.3635 under the trimf. Hence, we cannot identify a single MF which works well for all the three variants. It can be deduced that trimf and gbell MFs are better as in majority cases they yield lower RMSE.

5.1.1 Comparison of single-lexicon-based ANFIS with other approaches

Comparison of VADER lexicon-based ANFIS with VADER-specific approaches is shown in Table 7. It has been compared with two approaches—fuzzy rule approach (Vashishtha and Susan 2019) and the non-fuzzy approach (Hutto and Gilbert 2014). It can be clearly observed the our ANFIS outperforms all methods for all datasets. For instance, in Apple dataset RMSE of ANFIS is 0.45668, while other methods have 0.6127 and 0.6223. Table 8 presents the achievement of AFINN lexicon-based ANFIS with AFINN-specific methods. It has been compared with two approaches—fuzzy rule approach (Vashishtha and Susan 2019) and the non-fuzzy approach (Nielsen 2011). Our ANFIS has the lowest RMSE as compared to both approaches for all datasets. For SemEval 2017 dataset, ANFIS scored 0.47428, while others scored 0.6127 and 0.7046, respectively. Similar observations can be made for SentiWordNet lexicon-based ANFIS with SentiWordNet-specific approaches in Table 9. It has been compared with three approaches—fuzzy rule approach (Vashishtha and Susan 2019) and two non-fuzzy approaches: Cavalcanti (Cavalcanti et al. 2011) and Ortega (Ortega et al. 2013). Here also, ANFIS surpasses all methods, for example, for the SemEval 2016 dataset, RMSE for ANFIS is 0.45204 and for others: 0.554, 0.7164 and 0.8115. Hence, we can deduce that all single-lexicon-based ANFISs have lowest RMSE for all datasets compared to state-of-the-art SA models. But it is difficult to decide the best membership function (MF) for each lexicon among all datasets. In other words, we cannot conclude which is the best MF among trimf, gbellmf and gaussmf for the VADER, AFINN or SentiWordNet lexicons. In VADER, there is a confusion among trimf or gaussmf; in AFINN, we can choose gbellmf as it performs slightly better, while in SentiWordNet it is hard to choose between trimf or gbellmf. The execution

Table 7 Comparison of VADER lexicon-based ANFIS with VADER-specific methods

Dataset	Methods	RMSE
Apple	VADER SA (Hutto and Gilbert 2014)	0.6127
	Fuzzy rule (Vashishtha and Susan 2019a)	0.6223
	Vader-ANFIS	**0.45668**
Nuclear	VADER SA (Hutto and Gilbert 2014)	0.8013
	Fuzzy rule (Vashishtha and Susan 2019a)	0.4168
	Vader-ANFIS	**0.33174**
Sanders	VADER SA (Hutto and Gilbert 2014)	0.6353
	Fuzzy rule (Vashishtha and Susan 2019a)	0.5593
	Vader-ANFIS	**0.43453**
SemEval (2017)	VADER SA (Hutto and Gilbert 2014)	0.6137
	Fuzzy rule (Vashishtha and Susan 2019a)	0.6892
	Vader-ANFIS	**0.47814**
SemEval (2016)	VADER SA (Hutto and Gilbert 2014)	0.6309
	Fuzzy rule (Vashishtha and Susan 2019a)	0.8078
	Vader-ANFIS	**0.45195**
SemEval (2015)	VADER SA (Hutto and Gilbert 2014)	0.5764
	Fuzzy rule (Vashishtha and Susan 2019a)	0.689
	Vader-ANFIS	**0.45257**
STS (Go et al. 2009)	VADER SA (Hutto and Gilbert 2014)	0.4251
	Fuzzy Rule (Vashishtha and Susan 2019a)	0.7525
	Vader-ANFIS	**0.3686**
Airline	VADER SA (Hutto and Gilbert 2014)	0.5012
	Fuzzy rule (Vashishtha and Susan 2019a)	0.8552
	Vader-ANFIS	**0.388**
Twitter	VADER SA (Hutto and Gilbert 2014)	0.5296
	Fuzzy rule (Vashishtha and Susan 2019a)	0.7949
	Vader-ANFIS	**0.4336**
Reddit	VADER SA (Hutto and Gilbert 2014)	0.4648
	Fuzzy rule (Vashishtha and Susan 2019a)	0.7765
	Vader-ANFIS	**0.3918**

Bold values are the lowest RMSE values for each dataset

time for single-lexicon-based ANFIS is about an average of 10–15 min.

5.2 Multiple-lexicon-based ANFIS: MultiLexANFIS

To solve the above confusion of choosing the best membership function (MF) for a given lexicon, the MultiLexANFIS architecture has been proposed. This system integrates all the lexicons into a single unit. Using each lexicon, two scores: TweetPos, positive score, and TweetNeg, negative score, are computed as discussed in Sect. 3. Hence, this integrated system has a total of six inputs, with two MFs for each input, 64 rules and one sentiment label as output.

Table 10 displays the performance of MultiLexANFIS for all datasets. This system has been compared with single-lexicon-based ANFIS as well; the best (lowest) RMSE—combination of lexicon–membership function for each dataset, is represented in the second column of Table 10. For instance, for the Apple dataset, among single-lexicon ANFIS, VADER variant has achieved the lowest RMSE of 0.457 using the gbell MF, while MultiLexANFIS has attained 0.448 RMSE using the gauss MF, which is the lowest value. Similar analysis is accomplished for all the datasets. This integrated MultiLexANFIS achieves lower RMSE as compared to the single-lexicon-based ANFIS for almost all datasets, though for some datasets the results are comparable. We also note down that trimf: triangular membership function performs the best in MultiLexANFIS (7 out of 10 datasets) and gaussmf: Gaussian membership function is at second position. Thus, the problem of confusion among the MFs for all datasets has been resolved. Comparison amidst the single-lexicon

Table 8 Comparison of AFINN lexicon-based ANFIS with AFINN-specific methods

Dataset	Methods	RMSE
Apple	AFINN SA (Nielsen 2011)	0.6112
	Fuzzy rule (Vashishtha and Susan 2019a)	0.6402
	AFINN-ANFIS	**0.46623**
Nuclear	AFINN SA (Nielsen 2011)	0.846
	Fuzzy rule (Vashishtha and Susan 2019a)	0.4413
	AFINN-ANFIS	**0.39682**
Sanders	AFINN SA (Nielsen 2011)	0.6235
	Fuzzy rule (Vashishtha and Susan 2019a)	0.5634
	AFINN-ANFIS	**0.44414**
SemEval (2017)	AFINN SA (Nielsen 2011)	0.6127
	Fuzzy rule (Vashishtha and Susan 2019a)	0.7046
	AFINN-ANFIS	**0.47428**
SemEval (2016)	AFINN SA (Nielsen 2011)	0.6591
	Fuzzy rule (Vashishtha and Susan 2019a)	0.8303
	AFINN-ANFIS	**0.45331**
SemEval (2015)	AFINN SA (Nielsen 2011)	0.5649
	Fuzzy rule (Vashishtha and Susan 2019a)	0.718
	AFINN-ANFIS	**0.44535**
STS (Go et al. 2009)	AFINN SA (Nielsen 2011)	0.4321
	Fuzzy rule (Vashishtha and Susan 2019a)	0.8287
	AFINN-ANFIS	**0.38946**
Airline	AFINN SA (Nielsen 2011)	0.532
	Fuzzy rule (Vashishtha and Susan 2019a)	0.8819
	AFINN-ANFIS	**0.3873**
Twitter	AFINN SA (Nielsen 2011)	0.5444
	Fuzzy rule (Vashishtha and Susan 2019a)	0.8106
	AFINN-ANFIS	**0.4269**
Reddit	AFINN SA (Nielsen 2011)	0.483
	Fuzzy rule (Vashishtha and Susan 2019a)	0.8001
	AFINN-ANFIS	**0.4271**

Bold values are the lowest RMSE values for each dataset

ANFISs reveals that VADER lexicon performs best. This is observed in the second column of Table 10; among ten datasets, VADER lexicon has the lowest RMSE for six datasets. The execution time of MultiLexANFIS is about 30 min. The training error graph for MultiLexANFIS for the Apple dataset with gaussmf is shown in Fig. 4. The computed test error, 0.448, is depicted in Fig. 5.

5.2.1 Comparison of multiple-lexicon-based ANFIS: MultiLexANFIS with state of the art

The comparative analysis of our proposed MultiLexANFIS with eight state-of-the-art methods that include various fuzzy, non-fuzzy and deep learning models is presented in Table 11. The following are the state-of-the-art methods: (1) a deep learning model—recurrent neural network (RNN) developed by Nemes and Kiss (2021) for classifying tweets, (2) a deep learning model—long short-term memory (LSTM) model developed by Rahman et al. (2016) for text classification, (3) an n-gram model with SVM classifier developed by *c*, (4) an unsupervised fuzzy rule-based system for sentiment analysis which is a published previous work of authors (Vashishtha and Susan 2019a), (5) VADER SA by Hutto and Gilbert (2014), (6) AFINN SA by Nielsen (2011) approach, (7) SA using SentiWordNet lexicon by Cavalcanti et al. (2011) approach and (8) SA using SentiWordNet lexicon by Ortega et al. (2013) approach. The deep learning models—RNN and LSTM, are given the same input features that are fed to MultiLexANFIS. The models are built using Keras and Tensorflow tools; RNN model consists of embedding layer, followed by bidirectional RNN wrapper, dropout layer and dense layer, while LSTM model contains LSTM layer instead of bidirectional RNN layer. The error produced by

Table 9 Comparison of SENTIWORDNET lexicon-based ANFIS with SENTIWORDNET-specific methods

Dataset	Methods	RMSE
Apple	Cavalcanti (Cavalcanti et al. 2011)	0.7437
	Ortega (Ortega et al. 2013)	0.6449
	Fuzzy rule (Vashishtha and Susan 2019a)	0.655
	SENTIWORDNET-ANFIS	**0.48503**
Nuclear	Cavalcanti (Cavalcanti et al. 2011)	0.9205
	Ortega (Ortega et al. 2013)	0.8553
	Fuzzy rule (Vashishtha and Susan 2019a)	0.4168
	SENTIWORDNET-ANFIS	**0.36351**
Sanders	Cavalcanti (Cavalcanti et al. 2011)	0.795
	Ortega (Ortega et al. 2013)	0.6202
	Fuzzy rule (Vashishtha and Susan 2019a)	0.5655
	SENTIWORDNET-ANFIS	**0.45318**
SemEval (2017)	Cavalcanti (Cavalcanti et al. 2011)	0.6892
	Ortega (Ortega et al. 2013)	0.6693
	Fuzzy rule (Vashishtha and Susan 2019a)	0.7176
	SENTIWORDNET-ANFIS	**0.49243**
SemEval (2016)	Cavalcanti (Cavalcanti et al. 2011)	0.554
	Ortega (Ortega et al. 2013)	0.7164
	Fuzzy rule (Vashishtha and Susan 2019a)	0.8115
	SENTIWORDNET-ANFIS	**0.45204**
SemEval (2015)	Cavalcanti (Cavalcanti et al. 2011)	0.6839
	Ortega (Ortega et al. 2013)	0.6619
	Fuzzy rule (Vashishtha and Susan 2019a)	0.7222
	SENTIWORDNET-ANFIS	**0.4844**
STS (Go et al. 2009)	Cavalcanti (Cavalcanti et al. 2011)	0.5226
	Ortega (Ortega et al. 2013)	0.6555
	Fuzzy rule (Vashishtha and Susan 2019a)	0.7418
	SENTIWORDNET-ANFIS	**0.42304**
Airline	Cavalcanti (Cavalcanti et al. 2011)	0.4629
	Ortega (Ortega et al. 2013)	0.6804
	Fuzzy rule (Vashishtha and Susan 2019a)	0.8552
	SENTIWORDNET-ANFIS	**0.3961**
Twitter	Cavalcanti (Cavalcanti et al. 2011)	0.5654
	Ortega (Ortega et al. 2013)	0.7703
	Fuzzy rule (Vashishtha and Susan 2019a)	0.7949
	SENTIWORDNET-ANFIS	**0.4321**
Reddit	Cavalcanti (Cavalcanti et al. 2011)	0.5224
	Ortega (Ortega et al. 2013)	0.7256
	Fuzzy rule (Vashishtha and Susan 2019a)	0.7765
	SENTIWORDNET-ANFIS	**0.4502**

Bold values are the lowest RMSE values for each dataset

these models is higher than MultiLexANFIS, and this is because the parameter tuning and optimization in ANFIS are better than the tuning of weights during backpropagation in deep learning models. In Table 11, for instance we can observe for the Airline Twitter dataset, MultiLexANFIS has the lowest RMSE value—0.3843, while RNN has 0.4417 RMSE and LSTM has 0.4383 RMSE. Overall, we can deduce that after MultiLexANFIS, RNN is second best followed by LSTM. The third approach (Tripathy et al. 2016) creates different n-grams—unigrams, bigrams and trigrams from text, and assigns term frequency-inverse document frequency (TF-IDF) values to each n-gram; these numerical vectors are sent as input to SVM classifier. The errors produced by the n-gram model

Table 10 MultiLexANFIS: multiple-lexicon-based ANFIS

Datasets	Least among Lex Lex MF RMSE	TRIMF RMSE	GBELL RMSE	GAUSS RMSE
Apple	V Gbell— 0.457	0.475	0.452	**0.448**
Nuclear	V Gauss—0.332	0.374	0.356	**0.320**
Sanders	V Trimf—0.435	**0.440**	0.494	0.519
SemEval (2017)	A Trimf—0.474	**0.473**	0.474	0.473
SemEval (2016)	V Trimf—0.452	**0.455**	0.456	**0.454**
SemEval (2015)	A Gauss—0.445	**0.433**	0.447	0.445
STS (Go et al. 2009)	V Gauss—0.369	0.380	**0.366**	0.376
Airline	A Gbell—0.3873	**0.3843**	0.3860	0.3859
Twitter	A Gbell—0.4269	**0.4082**	0.4159	0.4127
Reddit	V Gbell—0.4036	**0.3918**	0.4104	0.4377

Bold values are the lowest RMSE value for each dataset. V stands for VADER and A stands for AFINN

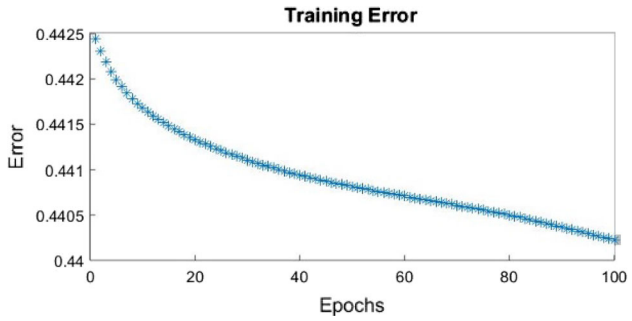


Fig. 4 Training error for Apple dataset with gaussmf in MultiLexANFIS

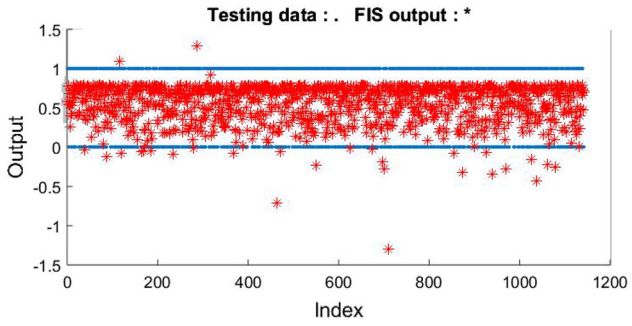


Fig. 5 Test error for Apple dataset with gaussmf in MultiLexANFIS (*red dots indicate the predicted values for testing data*) (color figure online)

are higher than MultiLexANFIS for all datasets, and this is because of missing fuzzy factor in the extracted features and in the model. The selected n-grams are based on frequency, while in our model we use the polarity and sentiment score of each word derived from lexicons; this way we are able to deal with natural language in a better way for SA. The proposed MultiLexANFIS is also compared to fuzzy rule classifier's (Vashishtha and Susan 2019a) three variants—VADER, AFINN and SentiWordNet. Our neuro-fuzzy model yields better results by using scores from sentiment lexicon as features for subjectivity classification (neutral non-neutral) compared to the unsupervised fuzzy rule classifier which uses the same scores for polarity classification (positive, negative or neutral). In fuzzy rule classifier, triangular (trimf) fuzzy membership is used and its parameters are fixed, while in our ANFIS model these parameters are tuned during each epoch, thereby producing lower errors. The next four state-of-the-art methods are unsupervised and non-fuzzy algorithms. They produce very high errors compared to all other approaches, and this is because they are simple and naïve in nature. Finally, we can conclude that the proposed MultiLexANFIS produces lower error for all datasets compared to other approaches and also it takes less execution time in contrast to single-lexicon-based ANFIS.

6 Conclusion

In this paper, a novel MultiLexANFIS has been proposed for social sentiment analysis. This architecture is an adaptive neuro-fuzzy inference system (ANFIS) that incorporates inputs from VADER, AFINN and SentiWordNet lexicons to perform sentiment analysis of social media content. The key highlights of this work are: (1) a novel neuro-fuzzy system: MultiLexANFIS that combines sentiment scores from multiple lexicons to classify tweets into neutral and non-neutral tweets, (2) a novel set of 64 rules for Sugeno-type fuzzy inference system—MultiLexANFIS, (3) single-lexicon-based ANFIS variants to classify tweets and (4) comparison of fuzzy approaches, non-fuzzy approaches and deep learning models with MultiLexANFIS that affirms the supremacy of our proposed neuro-fuzzy system. ANFIS has deployed optimization with the help of a hybrid learning algorithm; that uses least square method and gradient descent to tune parameters. We can deduce that MultiLexANFIS has the

Table 11 Comparison of MultiLexANFIS with the state of the art

Datasets	Least among Single-Lex ANFIS Lex MF RMSE	MultiLexANFIS RMSE	RNN (Nemes and Kiss 2021) RMSE	LSTM (Rahman et al. 2016) RMSE	n-gram SVM (Tripathy et al. 2016) RMSE	VADER—fuzzy rule (Vashishtha and Susan 2019a) RMSE	AFINN—fuzzy rule (Vashishtha and Susan 2019a) RMSE	SWN—fuzzy rule (Vashishtha and Susan 2019a) RMSE	VADER SA (Hutto and Gilbert 2014) RMSE	AFINN SA (Nielsen 2011) RMSE	SWN SA (Cavalcanti et al. 2011) RMSE	SWN SA (Ortega et al. 2013) RMSE
Apple	V Gbell-0.457	**0.448**	0.5176	0.5309	0.6689	0.6223	0.6402	0.655	0.6127	0.6112	0.7437	0.6449
Nuclear	V Gauss—0.332	**0.320**	0.4699	0.456	0.3245	0.4168	0.4413	0.4168	0.8013	0.846	0.9205	0.8553
Sanders	V Trimf—0.435	**0.440**	0.4856	0.4911	0.5527	0.5593	0.5634	0.5655	0.6353	0.6235	0.795	0.6202
SemEval (2017)	A Trimf—0.474	**0.473**	0.514	0.5254	0.6978	0.6892	0.7046	0.7176	0.6137	0.6127	0.6892	0.6693
SemEval (2016)	V Trimf—0.452	**0.454**	0.4632	0.4661	0.5462	0.8078	0.8303	0.8115	0.6309	0.6591	0.554	0.7164
SemEval (2015)	A Gauss—0.445	**0.433**	0.5615	0.5452	0.6909	0.6890	0.7180	0.7222	0.5764	0.5649	0.6839	0.6619
STS (Go et al. 2009)	V Gauss—0.369	**0.366**	0.5262	0.5244	0.5416	0.7525	0.8287	0.7418	0.4251	0.4321	0.5226	0.6555
Airline	A Gbell—0.3873	**0.3843**	0.4417	0.4383	0.4569	0.8552	0.8819	0.8816	0.5012	0.5320	0.4629	0.6804
Twitter_2019	A Gbell—0.4269	**0.4082**	0.5241	0.5275	0.5629	0.7949	0.8106	0.794	0.5296	0.5444	0.5654	0.7703
Reddit	V Gbell—0.4036	**0.3918**	0.5662	0.5627	0.5723	0.7765	0.8001	0.762	0.4648	0.483	0.5224	0.7256

Bold values are the lowest RMSE value for each dataset. V stands for VADER and A stands for AFINN

lowest RMSE for the ten benchmark datasets, as compared to other approaches. Results reveal that among the single-lexicon ANFISs, VADER lexicon performs best. The best membership in MultiLexANFIS is triangular (trimf), and Gaussian membership function (gaussmf) is at second position. The problem of confusion among the MFs for all datasets has been resolved using MultiLexANFIS. The MultiLexANFIS is computationally faster than the total combination of three single-lexicon-based ANFIS for all lexicons. Our neuro-fuzzy network is targeted for applications that aim to search neutral social media content posted by the users. Our results thus indicate an effective and computationally feasible solution to social sentiment analysis that incorporates multiple lexicons onto the same platform. These proposed novel rules are domain independent. We can apply these rules for any textual data that employs lexicons to extract positive and negative score for each textual document or sentence or review. The proposed system can handle only those twitter datasets which contain neutral tweets. At present, it cannot deal with sarcasm and mixed tweets, i.e., the tweets that contain both positive and negative opinions. In future, we will compute the sentiment of mixed tweets by enhancing our model using advanced fuzzy deep neural networks.

Funding This study was not funded by any agency.

Declarations

Conflict of interest Both authors are associated with the Department of Information Technology, Delhi Technological University.

Ethical approval This article does not contain any studies with animals performed by any of the authors.

Human and animal rights This article does not contain any studies with human participants or animals performed by any of the authors.

References

Acampora G, Cosma G (2014) A hybrid computational intelligence approach for efficiently evaluating customer sentiments in E-commerce reviews. In: 2014 IEEE symposium on intelligent agents (IA). IEEE, pp 73–80

Airline Twitter Dataset (2019) https://www.kaggle.com/crowdflower/twitter-airline-sentiment. Accessed on 29 Aug 2021

Ahuja R, Rastogi H, Choudhuri A, Garg B (2015) Stock market forecast using sentiment analysis. In: IEEE 2nd international conference on computers for sustainable global development, pp 1008–1010

Anagha M, Kumar RR, Sreetha K, Raj PR (2015) Fuzzy logic based hybrid approach for sentiment analysis of Malayalam movie reviews. In: 2015 IEEE international conference on signal processing, informatics, communication and energy systems (SPICES). IEEE, pp 1–4

Apple Twitter Dataset (2019). https://data.world/crowdflower/apple-twitter-sentiment. Accessed 29 Aug 2021

Asimuzzaman M, Nath PD, Hossain F, Hossain A, Rahman RM (2017) Sentiment analysis of bangla microblogs using adaptive neuro fuzzy system. In: 2017 13th international conference on natural computation, fuzzy systems and knowledge discovery (ICNC-FSKD). IEEE, pp 1631–1638

Baccianella S, Esuli A, Sebastiani F (2010) Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. Lrec 10(2010):2200–2204

Banerjee S, Pedersen T (2002) An adapted Lesk algorithm for word sense disambiguation using WordNet. In: International conference on intelligent text processing and computational linguistics. Springer, Berlin, Heidelberg, pp 136–145

Çakıt E, Karwowski W, Servi L (2019) Application of soft computing techniques for estimating emotional states expressed in Twitter® time series data. Neural Comput Appl 1–14:2019

Cavalcanti DC, Prudêncio RBC, Pradhan SS, Shah JY, Pietrobon RS (2011) Good to be bad? Distinguishing between positive and negative citations in scientific impact. In: 2011 IEEE 23rd international conference on tools with artificial intelligence. IEEE, pp 156–162

Cosma G, Acampora G (2016) Neuro-fuzzy sentiment analysis for customer review rating prediction. In: Sentiment analysis and ontology engineering. Springer, Cham, pp 379–397

Ghiassi M, Skinner J, Zimbra D (2013) Twitter brand sentiment analysis: a hybrid system using n-gram analysis and dynamic artificial neural network. Expert Syst Appl 40(16):6266–6282

Guerra PHC, Veloso A, Meira Jr W, AlmeidaV (2011) From bias to opinion: a transfer-learning approach to real-time sentiment analysis. In: Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, pp 150–158

Gupta N, Sharma G, Sharma RS (2014) A comparative study of ANFIS membership function to predict ERP user satisfaction using ANN and MLRA. Int J Comput Appl 105(5)

Go A, Bhayani R, Huang L (2009) Twitter sentiment classification using distant supervision. CS224N Project Report, Stanford 1, no. 12 (2009)

Hutto CJ, Gilbert E (2014) Vader: a parsimonious rule-based model for sentiment analysis of social media text. In: Eighth international AAAI conference on weblogs and social media

Indhuja K, Reghu RP (2014) Fuzzy logic based sentiment analysis of product review documents. In: 2014 first international conference on computational systems and communications (ICCSC). IEEE, pp 18–22

Jang JSR (1993) ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 23(3):665–685

Jang JSR, Sun CT, Mizutani E (1997) Neuro-fuzzy and soft computing-a computational approach to learning and machine intelligence [Book Review]. IEEE Trans Autom Control 42(10):1482–1484

Jefferson C, Liu H, Cocea M (2017) Fuzzy approach for sentiment analysis. In: 2017 IEEE international conference on fuzzy systems (FUZZ-IEEE). IEEE, pp 1–6

Kamil A, Rustamov S, Clements MA, Mustafayev E (2018) Adaptive neuro-fuzzy inference system for classification of texts. In: Recent developments and the new direction in soft-computing foundations and applications. Springer, Cham, pp 63–70

Kalarani P, Brunda SS (2019) Sentiment analysis by POS and joint sentiment topic features using SVM and ANN. Soft Comput 23(16):7067–7079

Katta P, Hegde NP (2019) A hybrid adaptive neuro-fuzzy interface and support vector machine based sentiment analysis on political twitter data. Int J Intell Eng Syst 12(1):165–173

Khan W, Malik U, Ghazanfar MA, Azam MA, Alyoubi KH, Alfakeeh AS (2020) Predicting stock market trends using machine learning algorithms via public sentiment and political situation analysis. Soft Comput 24(15):11019–11043

Khattak A, Paracha WT, Asghar MZ, Jillani N, Younis U, Saddozai FK, Hameed IA (2020) Fine-grained sentiment analysis for measuring customer satisfaction using an extended set of fuzzy linguistic Hegdes. Int J Comput Intell Syst 13(1):744–756

Kulkarni S, Vidhate A, Anuradha G (2014) Neuro fuzzy approach for predicting sales performance of movies considering sentiments in online reviews

Kulkarni S, Nikumbh PJ, Anuradha G, Nikam S (2012) S-ANFIS: sentiment aware adaptive network-based fuzzy inference system for Predicting Sales Performance using Blogs/Reviews. Int J Multidiscip Cryptol Inf Secur 1(2)

Koppel M, Schler J (2005) Using neutral examples for learning polarity. In: International joint conference on artificial intelligence. Lawrence Erlbaum Associates Ltd, vol 19, p 1616

Koppel M, Schler J (2006) The importance of neutral examples for learning sentiment. Comput Intell 22(2):100–109

Liu H, Cocea M (2017) Fuzzy rule based systems for interpretable sentiment analysis. In: 2017 ninth international conference on advanced computational intelligence (ICACI). IEEE, pp 129–136

Madhusudhanan S, Moorthi M (2019) Optimized fuzzy technique for enhancing sentiment analysis. Clust Comput 22(5):11929–11939

Mazumder P, Chowdhury NA, Bhuiya MA, Akash SH, Rahman RM (2018) A fuzzy logic approach to predict the popularity of a presidential candidate. In: Modern approaches for intelligent information and database systems. Springer, Cham, pp 63–74

Nagamanjula R, Pethalakshmi A (2020) A novel framework based on bi-objective optimization and LAN 2 FIS for Twitter sentiment analysis. Soc Netw Anal Min 10:1–16

Nemes L, Kiss A (2021) Social media sentiment analysis based on COVID-19. J Inf Telecommun 5(1):1–15

Nguyen TL, Kavuri S, Lee M (2019) A multimodal convolutional neuro-fuzzy network for emotion understanding of movie clips. Neural Netw 118:208–219

Nielsen FÅ (2011) Afinn.Richard Petersens Plads, Building 321 (2011)

Nuclear Twitter Dataset (2019) https://data.world/crowdflower/emotions-about-nuclear-energy. Accessed 29 Aug 2021

Ortega R, Fonseca A, Montoyo A (2013) SSA-UO: unsupervised Twitter sentiment analysis. In: Second joint conference on lexical and computational semantics (* SEM), vol 2, pp 501–507

Padmaja K, Hegde NP (2019) Twitter sentiment analysis using adaptive neuro-fuzzy inference system with genetic algorithm. In: 2019 3rd international conference on computing methodologies and communication (ICCMC). IEEE, pp 498–503

Parveen H, Pandey S (2016) Sentiment analysis on twitter data-set using naive Bayes algorithm. In: 2016 2nd international conference on applied and theoretical computing and communication technology (iCATccT). IEEE, pp 416–419

Phan HT, Nguyen NT, Van Cuong T, Hwang D (2019) A method for detecting and analyzing the sentiment of tweets containing fuzzy sentiment phrases. In: 2019 IEEE international symposium on innovations in intelligent systems and applications (INISTA). IEEE, pp 1–6

Rahman L, Mohammed N, Al Azad AK (2016) A new LSTM model by introducing biological cell state. In: 2016 3rd international conference on electrical engineering and information communication technology (ICEEICT). IEEE, pp 1–6

Rustamov S, Mustafayev E, Clements MA (2013) Sentiment analysis using Neuro-Fuzzy and Hidden Markov models of text. In: 2013 Proceedings of IEEE Southeastcon. IEEE, pp 1–6

Rustamov S, Mustafayev E, Clements M (2013) Sentence-level subjectivity detection using neuro-fuzzy models. In: Proceedings of the 4th workshop on computational approaches to subjectivity, sentiment and social media analysis, pp 108–114

Rustamov S (2018) A hybrid system for subjectivity analysis. Advances in Fuzzy Systems 2018

Sanders Twitter Dataset (2019) http://www.sananalytics.com/lab. Accessed 29 Aug 2021

Sasikala P, Sheela LMI (2020) Sentiment analysis of online product reviews using DLMNN and future prediction of online product using IANFIS. J Big Data 7(1):1–20

Siddiqua UA, Ahsan T, Chy AN (2016) Combining a rule-based classifier with weakly supervised learning for twitter sentiment analysis. In: 2016 international conference on innovations in science, engineering and technology (ICISET). IEEE, pp 1–4

SemEval (2017), http://alt.qcri.org/semeval2017/task4/. Accessed 29 Aug 2021

SemEval (2016), http://alt.qcri.org/semeval2016/task4/. Accessed 29 Aug 2021

SemEval (2015), http://alt.qcri.org/semeval2015/task10/. Accessed 29 Aug 2021

Serrano-Guerrero J, Romero FP, Olivas JA (2021) Fuzzy logic applied to opinion mining: a review. Knowl-Based Syst:107018

Sugeno M, Kang GT (1988) Structure identification of fuzzy model. Fuzzy Sets Syst 28(1):15–33

Susan S, Khowal SK, Kumar A, Kumar A, Yadav AS (2013) Fuzzy min-max neural networks for business intelligence. In: 2013 international symposium on computational and business intelligence. IEEE, pp 115–118

Talpur N, Salleh MNM, Hussain K (2017) An investigation of membership functions on performance of ANFIS for solving classification problems. In: IOP conference series: materials science and engineering. IOP Publishing, vol 226, no 1, p 012103

Tripathy A, Agrawal A, Rath SK (2016) Classification of sentiment reviews using n-gram machine learning approach. Expert Syst Appl 57:117–126

Twitter_2019 (2019) https://www.kaggle.com/cosmos98/twitter-and-reddit-sentimental-analysis-dataset. Accessed 29 Aug 2021

Vashishtha S, Susan S (2018) Fuzzy logic based dynamic plotting of mood swings from tweets. In: International conference on innovations in bio-inspired computing and applications, Springer, Cham, pp 129–139

Vashishtha S, Susan S (2019a) Fuzzy rule based unsupervised sentiment analysis from social media posts. Expert Syst Appl 138:112834

Vashishtha S, Susan S (2019b) Sentiment cognition from words shortlisted by fuzzy entropy. IEEE Trans Cogn Dev Syst. https://doi.org/10.1109/TCDS.2019.2937796

Vashishtha S, Susan S (2020a) Inferring sentiments from supervised classification of text and speech cues using fuzzy rules. Procedia Comput Sci 167:1370–1379

Vashishtha S, Susan S (2020b) Fuzzy interpretation of word polarity scores for unsupervised sentiment analysis. In: 2020 11th international conference on computing, communication and networking technologies (ICCCNT). IEEE (in press)

Vashishtha S, Susan S (2021) Highlighting keyphrases using senti-scoring and fuzzy entropy for unsupervised sentiment analysis. Expert Syst Appl 169:114323

Valdivia A, Luzíón MV, Herrera F (2017) Neutrality in the sentiment analysis problem based on fuzzy majority. In: 2017 IEEE international conference on fuzzy systems (FUZZ-IEEE). IEEE, pp 1–6

Valdivia A, Luzón MV, Cambria E, Herrera F (2018) Consensus vote models for detecting and filtering neutrality in sentiment analysis. Inf Fusion 44(2018):126–135

Wu KY, Zhou MC, Lu XS, Huang L (2017) A fuzzy logic-based text classification method for social media data. In: 2017 IEEE international conference on systems, man, and cybernetics (SMC). IEEE, pp 1942–1947

Xie X, Ge S, Hu F, Xie M, Jiang N (2019) An improved algorithm for sentiment analysis based on maximum entropy. Soft Comput 23(2):599–611

Yakubu H, Kwong CK, Lee CK (2021) A multigene genetic programming-based fuzzy regression approach for modelling customer satisfaction based on online reviews. Soft Comput 25(7):5395–5410

Žunić E, Djedović A, Avdagić Z (2016) Decision support system for candidates classification in the employment process based on ANFIS method. In: 2016 XI International symposium on telecommunications (BIHTEL). IEEE, pp 1–6

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

ARTICLE IN PRESS

Materials Today: Proceedings xxx (xxxx) xxx

Contents lists available at ScienceDirect

Materials Today: Proceedings

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

Performance analysis of inclined microjet impingement heat sink with porous medium

Harshit Jain *, Kanishk Yadav, Kapeesh Kumar, M. Zunaid

Department of Mechanical Engineering, Delhi Technological University, Delhi, India

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Microjet array
Heat sink
Inclined microjets
Electronics cooling
Pressure drop
Porous medium

ABSTRACT

The purpose of this research is to find out how well inclined microjet array heat sinks with porous pillars in microchannels perform in electronic cooling applications. The investigation is carried out through CFD analysis for laminar flow with water as heat transfer fluid. Outlet temperature and pressure drop variations are analysed by varying the number of nozzles, jet spacing and porosity of the porous pillars for determining the thermal and hydraulic performance of the inclined micro jet heat sink. The performance of the microjet heat sink was evaluated for a jet array consisting of 2, 6, 10, 14. 18 jets with porous pillars of different porosity values. The idea behind adding porous medium in microchannels was to increase the heat transfer rate between the solid–liquid interface. A constant heat flux of 200 kW/m^2 was provided at the bottom surface of the copper substrate to simulate the actual heat generation in electronic devices. The results showed there was a significant improvement in thermal effectiveness in the inclined microjet heat sink with the addition of porous medium. The outlet temperature of water increased by 4.5 °C with the addition of porous medium. Although some rise in pressure drop was also observed, the penalty of it was not as significant when compared with the improvement in heat transfer performance.
Copyright © 2022 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the First International Conference on Design and Materials (ICDM)-2021

1. Introduction

Electronic device advancements, such as integrated circuit microminiaturization, have allowed for the insertion of more components into a given surface area. As a result, heat output by power electrical gadgets has increased. This has led to the development of new cooling techniques such as microchannel heat sink, sprays and impinging jets to improve the heat dissipation mechanism. Electronic device temperature regulation is of utmost importance as it maintains the electronic device within the normal operating temperature and prolongs its life.

Microjet heat sink devices provide efficient solution of high heat flux from a smaller surface area of electronic devices. The microchannel heat sink devices were initially studied by Tuckerman and Pease [1]. They found that liquid flows from inlet to outlet has a lower temperature at inlet and higher temperature at outlet. The microchannel plate's and individual microchannels' geometric arrangement has a significant impact on single-phase temperature distribution, according to X.F. Peng et al. [2], and the impact on laminar and turbulent convection differs dramatically. T.M. Harms et al. [3] conducted a similar study and discovered that in terms of flow and heat transfer, a microchannel system for creating laminar flow beats a comparable single-channel system for turbulent flow. Furthermore, narrowing the channel and increasing the microchannel depth increase flow and heat transfer, according to the research.

C. P. Tso et al. [4] analysed the Brinkman number (Br) and found that Br determines the basic limit for microchannel dimension reduction as well as a measure of the relative relevance of viscous dissipation in heat sink. In the laminar regime, it should be more significant than in the transition and turbulent regimes. T.M. Adams et al. [5] simulated for turbulent flow in circular microchannels. According to the statistics of experimental result data, when the channel width reduces and the Reynolds number rises, the amount of amplification increases.

C. Yang et al. [6] investigation results clearly revealed that the presence of an electric double layer field considerably influences the liquid flow and heat transfer in heat sink for aqueous solutions with low ionic concentrations. According to D. Y Lee et al. [7] and A. G. Fedorov et al. [8], the thermal performance of jet impingement

* Corresponding author.
E-mail address: harshitjain8991@gmail.com (H. Jain).

https://doi.org/10.1016/j.matpr.2021.12.066
2214-7853/Copyright © 2022 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the First International Conference on Design and Materials (ICDM)-2021

Please cite this article as: H. Jain, K. Yadav, K. Kumar et al., Performance analysis of inclined microjet impingement heat sink with porous medium, Materials Today: Proceedings, https://doi.org/10.1016/j.matpr.2021.12.066

Nomenclature

As	cross-sectional area (mm^2)
dn	microjet diameter (mm)
Hc	height of the liquid channel (mm)
Lj	vertical height of the microjet nozzle (mm)
Ly	y directional length (mm)
Δp	pressure drop (kPa)
ΔT	temperature rise (K)
To	liquid outlet temperature (K)
α	design variable (Hc/dn)
Sn	distance of nozzle from the edge
φ	Porosity
V_R	Viscous resistance
K_p	Permeablity
∇	Gradient Operator
V	Velocity of the fluid
μ	Dynamic viscosity of fluid
K	Thermal conductivity of fluid
CFD	computational fluid dynamics
Dp	diameter of porous pillar (mm)
j	number of jets
Lx	x directional length (mm)
Lz	z directional length (mm)
Dn	inter jet spacing (mm)
Tavg	average temperature of solid–liquid interface (K)
Ts	thickness of the copper substrate base (mm)
3-D	three dimensional
Sp	distance between two pillars
ρ	density
C_2	Inertial resistance
R_e	Reynolds number
∇^2	Laplacian operator
P	Pressure
C_p	Specific heat at constant pressure
Cf	Forchheimer constant

without any treatment of spent flow is much inferior than that of microchannel cooling, independent of the intended plate dimension. X. F. Peng et al. [9 10 11] discovered that for single-phase flow through microchannels, there is a liquid flow mode transition area beyond which the heat transfer coefficient is largely independent of the wall temperature. X. F. Peng et al. [10] discovered that flow mode and heat transfer regime transformations occur at a substantially lower Reynold number than in the normal rectangular cross-section case.

Weilin Qu et al. [12] investigated the heat transfer characteristics of water flow in silicon micro - channels and discovered that the numerically calculated Nusselt number in micro - channels is lower than the empirically observed Nusselt number. Myung et al. [13] investigated the thermal performance of two-phase cooling in a hybrid cooling system involving attributes of both jets and microchannel flow. They found that cooling effects were more pronounced for jet impingement as compared to microchannel flow. Shin et al. [14] used dielectric coolant HFE-7000 as heat transfer fluid and found that heat transfer increased with jet velocity with highest value of 72 kW/m^2 at Reynold number of 4057. Xunfeng et al. [15] used a combination of jet array impinging on micro pin-fins surface array for improvement in heat transfer performance and found that the area average Nusselt number varied as a function of wetted surface area and thermal conduction resistance from pin-fin material.

V.S Devahdhanush et al. [16] experimentally examined the effect on heat flux for round single jets and jet arrays for R-134a. High CHF was observed for an increase in jet velocity, nozzle diameter and saturation temperature. Cong Hiep et al. [17] experimental study of two-phase microchannel heat sink with fins was performed. Results showed that reduced subcooling resulted in 25% increase in heat transfer coefficient but with the penalty of increase in pressure drop. Yonghai Zhang et al. [18] showed that pin-finned surfaces had a better thermal performance than smooth surfaces and Reynold Number plays a very important role in deciding heat transfer performance of microchannel heat sink. Xing Ju et al. [19] developed a novel heat sink where the flow distribution was inspired by chessboard pattern which showed improvement in temperature uniformity. Seung Ho Han et al. [20] carried out research to find out the heat transfer effect of milli-scale flat, convex and concave surfaces. Results showed that Nusselt number is very sensitive to nozzle to surface distance. D.Jacob Butterfield et al. [21] confirmed that using superhydrophobic (SH) surfaces possess anti fouling properties but also reduce the heat transfer effect. Yung feng Cui et al. [22] investigated the phenomenon of bubble dynamics and pressure oscillations of jet array impingement boiling with subcooled water and found that nucleate boiling occurs near walls of jet collision. Ramaswamy Devakottai Sundaram et al. [23] showed that placing short height rough elements on the impinging wall results in increased heat dissipation without much penalty on piping power. R. Jenkins et al. [24] investigated two-phase jet array impingement heat sink and found that nozzle is insensitive for R_e of fully developed boiling.

Amir Gorodetsky et al. [25] investigated a hybrid heat sink comprised of a microchannel and a jet array. When compared to a traditional heat sink of identical dimensions, the jet array resulted in a 20 percent increase in heat transmission. A.J Robinson et al. [26] designed a microchannel heat sink using additive manufacturing and found its thermal performance to be extraordinary with heat transfer rate of 400 kW/m^2 at flow rate of 0.006 kg/s. Jeong Heon Marcus Vinicus P. Carnerio [27] performed analysis of multiple orifice spray cooling heat sink using an array of parallel and oblique jets that impinge with the heated surface. Hey found that the highest value was obtained for the oblique jet array.

Songkran Wirryasant et al. [28] studied different fin shape design for jet impingement heat sink. They found that circular shape fin to be most efficient among all fin shapes with thermal resistance 12% to 25% lower. P.Naphon et al. [29] used TiO_2 nanofluid as heat transfer fluid and They discovered that with a nanofluid concentration of 0.015 percent, the convective heat transfer coefficient improved by 18.56 percent. Prachant Singh [30] used cylindrical, cubicle, and concentric form components to investigate the influence of impinging surface roughness in jet impingement heat sinks. In terms of fin efficiency, the concentric roughness shape produced the best results.

Liang Gong et al. [31] employed a computational approach to demonstrate that the configuration of a microchannel heat sink is a crucial factor in heat transfer performance. H.C Cui et al X.T Lai et al [32] simulated straight ribs and oblique ribs jet impingement heat sink and observed that oblique rib achieved more uniform and lower overall temperature than straight rib. Dorin Lelea et al [33] observed that change in viscosity has a dominating effect on the thermal performance of a heat sink. Paisasn Naphon et al [34] performed an experiment on a mini rectangular fins jet heat

sink and found it to have lower overall temperature than conventional heat sink.

Manish Kumar et al. [35] analysed meshless methods such as Element Free Galerkin Method (EFGM) and Meshless Local Petro-Galerkin Method (MLPG) to solve elasto-statics problem. Meshless methods do not require discretisation of surface, which is otherwise required for Finite Element Method (FEM). The results obtained from meshless methods were found to converge faster and more accurately as compared to FEM.

The literature reviewed above revealed that a substantial amount of experimental and computational work has been done in the topic of microjet heat sinks. Because no substantial work has been documented in the open literature on the impact of porous material on inclined microjet heat sinks, the current study analyses it. A three-dimensional CFD technique is used to study the thermal performance.

Description of the Geometric Model

Copper-based inclined microjet impingement models are depicted in a schematic diagram in Fig. 1(a). The microchannel substrate has a surface area of 12 × 12 mm^2. The impingement heat sink plate has a thickness of 0.8 mm and is designed with numerous microjets. The height of the copper substrate's base is 0.1 mm, while the liquid body is 0.3 mm long. The performance of heat sink with 2, 6, 10, 14, and 18 jets (Fig. 1) is evaluated. Some geometric parameters were held constant throughout the analysis, such as the area and depth of the substrate, jet diameter, the height of the fluid channel, and the nozzle's length. A constant diameter of 0.1 mm was used for the study. At the top of the fluid channel, there are slanted jets. The jet's inclination to the impingement surface held steady at 45 degrees. To reach the outlet port, the spent fluid must flow via the fluid channel. To simulate chip heat creation in the solid substrate, heat flux of constant value is delivered at the module's base. The liquid from the microjet nozzles hits the solid heated copper substrate, then escapes through the heat sink after cooling the bottom surface (see Fig. 2).

2. Governing equation

The equations used for conversion of mass, momentum, energy, and conjugate heat transfer are given as

Continuity equation

$$\nabla.\rho \mathrm{V} = 0$$

Momentum equation

$$(V.\nabla)\rho \mathrm{V} = -\nabla \mathrm{P} + \mu \nabla^2 \mathrm{V}$$

Energy equation

$\rho C_p(\mathrm{V}.\nabla \mathrm{T}) = \nabla.(\mathrm{K}\nabla \mathrm{T})$ (For fluid)

Thermal equation is given by

$\nabla(\nabla.T) = 0$(For the substrate conduction)

$$\text{Porosity}(\Phi) = \frac{\text{volume occupied by the liquid}}{\text{total volume of the material}}$$

The Brinkman-Hazen-Dupit-Darcy equation was used to compute the form coefficient (FC) and viscous coefficient. [36]. C f is set to 0.55 and Kp is fixed as 10^{-7} [37]. Form coefficient (FC) is defined as

$$FC = (Cf)/(\sqrt{Kp})$$

$$V_R = \Phi/Kp$$

$$C_2 = 2 \times FC \times \Phi^2$$

3. Numerical model

Three-dimensional numerical analysis was carried out in ANSYS Fluent for solving Navier Stokes equations under steady flow laminar conditions. The hexahedral mesh was generated with a finer grid near the jet nozzles and porous pillars to account for high temperature and pressure gradient. A grid size of 1,200,000 was used for the simulation. The fluid channel's inlet and exit were given boundary conditions of mass flow input with mass flow rate of 1.22×10^{-4} kg/s and outlet aspressure Outlet. The bottom surface of the copper solid substrate was given a uniform heat flux of 200 kW/m^2 and all other walls were kept adiabatic. The walls of nozzles and channels were subjected to a no-slip condition (V = 0). The interface between solid and fluid was coupled to account for conjugate heat transfer between the two media. Porous medium was created in ANSYS by changing the cell zone condition of pillars as porous zone. Porosity, viscous resistance (1/m) and inertial resistance values were provided from the Table 1 to the porous pillars to accurately simulate the effect of porosity on heat transfer and pressure drop for the microjet heat sink.

The energy and continuity residuals converged well below the value of 10^{-6}. The analysis of temperature distribution across the surface of substrate was carried out in ANSYS by temperature contour diagrams and average surface temperature was evaluated using surface integral function. Pressure drop was calculated using difference in pressure of inlet and outlet using area weighted average function in ANSYS.

4. Results & analysis

The grid-independent test was carried out for the designed heat sink and a mesh of approximately 1,200,000 was found suitable for $\alpha = 3$ and a heat sink of dimensions 12 × 12 mm^2. For a grid of size more than 1,200,000, the error was less than 1% and for a grid of size less than 600,000, it was found to be more than 5%. Hence, in order to reduce calculation time while maintaining the correctness of the results, a grid size of 1,200,000 was chosen.

Further, Simulations of laminar flow inclined jet impingement heat sinks by M. Zunaid et al. [38] were used to validate the current numerical approach. Fig. 3 shows that numerical results carried out in the present scheme match well with the results obtained by M. Zunaid et al. [38].

The numerical simulations were carried out for microjet heat sinks for 2, 6, 10, 14, 18 jets. The inclination of the microjet was kept at a 45-degree angle measured from the base of the fluid channel's top surface. The outlet of the microjet was kept in an opposite direction to the direction of the heat sink outlet so as to increase the area covered by the water for maximum heat transfer. The mass flow rate was kept uniform at 1.22×10^4 kg/s. The diameter of the nozzles and height of the fluid channel was kept constant as the effect of these variables has already been discussed in previous literature. M. Zunaid et al. [38] observed the highest thermal performance for design parameter = 3, hence the same value was used in the present study.

The novel idea in the present study design was to add a porous medium in the fluid channel so as to increase the heat transfer effect between fluid and solid substrate. Two designs were modelled to study the effect of porous medium on the thermal performance and pressure drop values of the heat sink by varying the porosity of the medium (φ). The first design consisted of 25 porous pillars of cylindrical shape situated inside the fluid domain in a symmetrical manner. The other design consisted of removing the

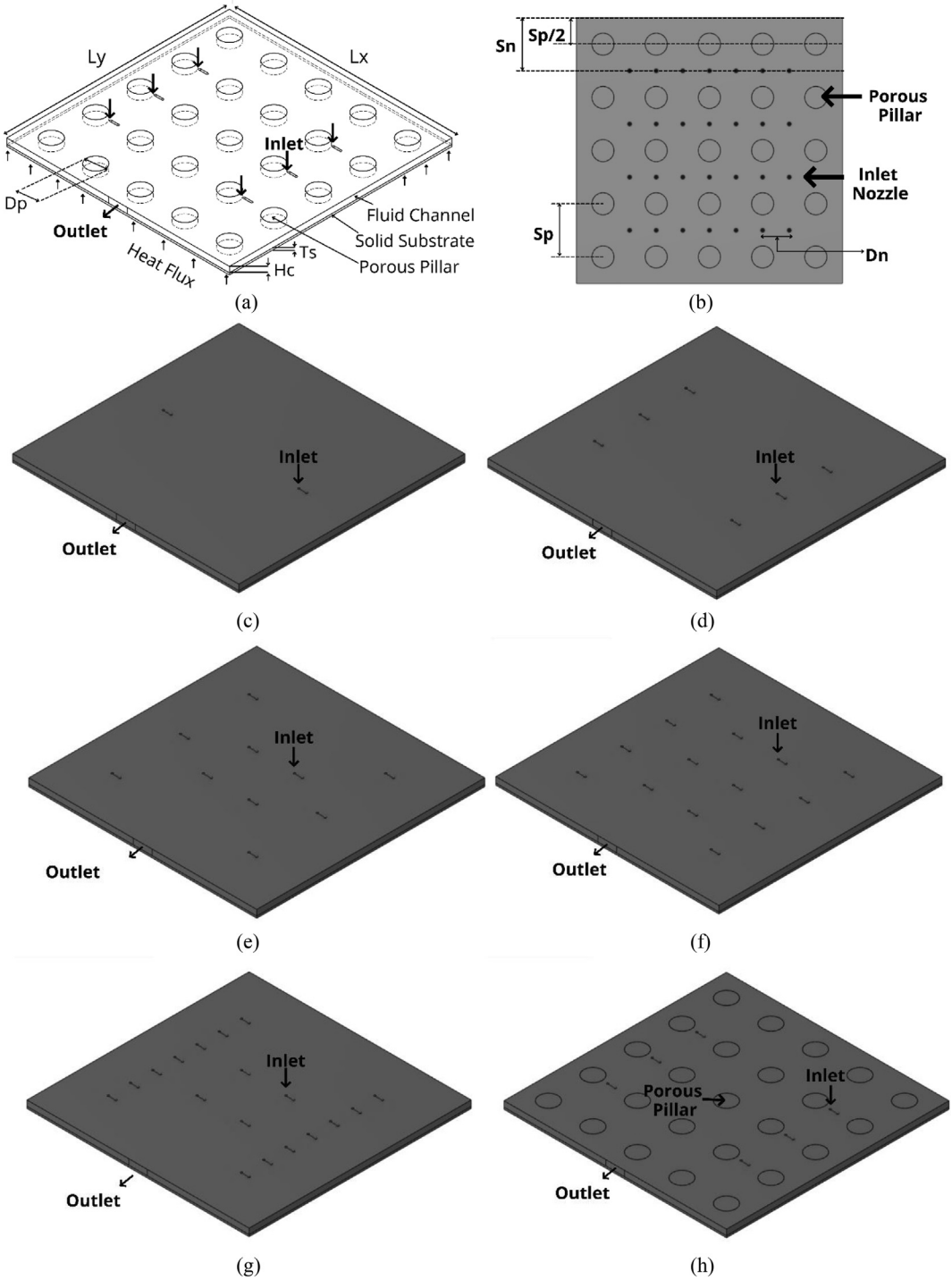


Fig. 1. (a) Isometric view & (b) top view of schematic illustration of a heat sink with (c) 2 nozzles, (d) 6 nozzles, (e) 10 nozzles, (f) 14 nozzles, (g) 18 nozzles, (h) 6 nozzles with porous pillars.

fluid domain altogether and replacing it with a porous medium. The effect of varying the porosity was studied with respect to outlet temperature of the fluid and pressure drop in the fluid channel. The outlet temperature of the fluid was considered as a performance parameter because the change in the temperature of the outlet is an indication of the amount of heat transfer between the solid–fluid interface.

Fig. 4(a) shows the variation of pressure drop with several nozzles for both with and without porous pillars in fluid channels. Fig. 4(a) shows that as the number of microjets increases, the pressure drop reduces for both porous and non-porous fluid channels. This is due to the fact that as the number of jets increases, the flow rate per jet reduces, resulting in a decrease in the pressure drop. It can also be observed that the pressure drops in the microchannel with porous pillars is more than the microchannel without porous pillars. This is a result of a higher-pressure gradient caused by the porous element. With an increase in the number of nozzles, however, the difference in pressure drop is not significant. The reason for this is because the flow rate per jet has decreased.

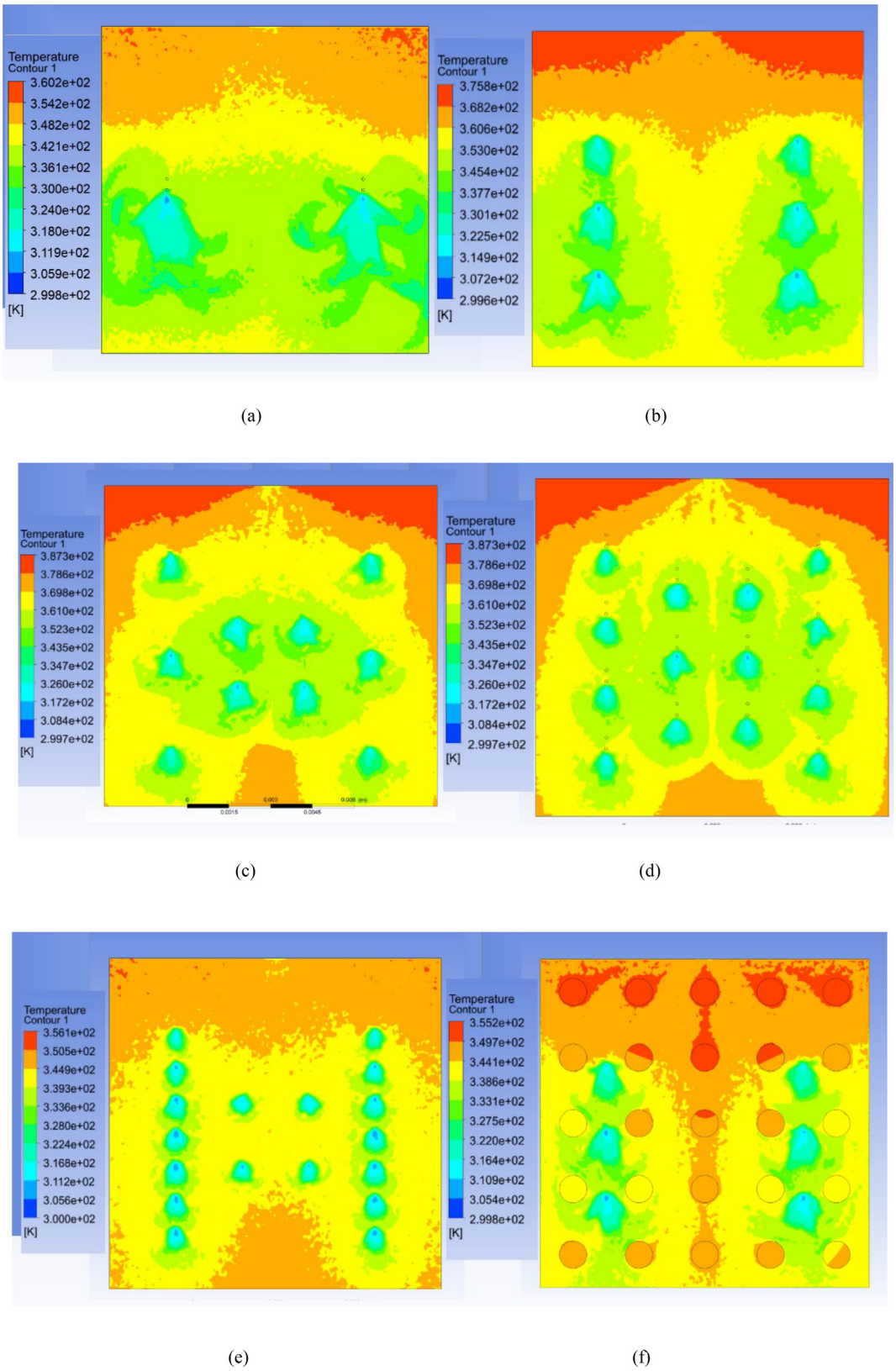


Fig. 2. Temperature distribution at the liquid–solid interface for (a) 2 nozzles, (b) 6 nozzles, (c) 10 nozzles,(d) 14 nozzles, (e)18 nozzles, (f) 6 nozzles with porous pillars.

Fig. 5(a) shows the variation of outlet temperature with change in the number of jets for both fluid domains with and without porous pillars. It is clearly visible from the figure that the outlet temperature of the water is higher for heat sink containing porous pillars as compared to the conventional heat sink for all the values of microjets. This provides evidence that microjet heat sinks containing porous pillars are more thermally efficient, as a higher value of outlet temperature means more heat transfer takes place

Table 1
Porosity Parameters.

Φ	V_R (1/m^2)	C_2 (1/m)
14%	14×10^5	73.5
22%	22×10^5	179.5
35%	35×10^5	426.17
50%	50×10^5	869.75
62%	62×10^5	1337.32
75%	75×10^5	1956.93
90%	90×10^5	2817.99

between the fluid and copper plate. It can also be seen that more

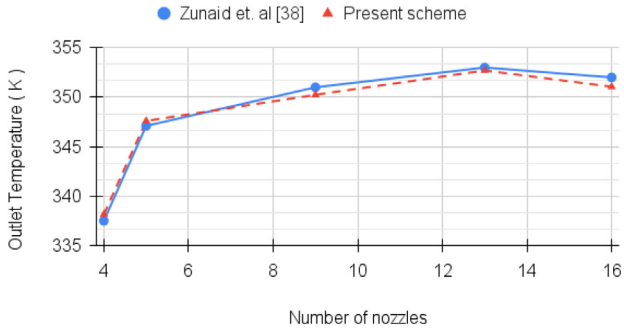


Fig. 3. Validation of numerical findings by comparing them to the results provided by M. Zunaid et al. [38].

number nozzles result in a higher value of outlet temperature. Hence using more nozzles combined with porous pillars can result in a better performing heat sink.

Fig. 4(b) and 5 (b) are for microjet heat sink with the fluid channel, fully replaced by porous media for the number of inclined nozzles equal to 6. Fig. 5(b) show the variation of outlet temperature for fully porous and no porous water channel. It can be clearly seen from the figure that the outlet temperature is higher for heat sink containing porous media by about 4.5 degrees Celsius. Hence it can be concluded that the addition of porous media in the microchannel increases the thermal performance of the heat sink. However, it must be noted that the variation in the outlet temperature with change in the porosity is almost negligible.

Fig. 4(b) shows the change in pressure drop for change in the porosity of the porous medium. The Figure shows that the pressure is higher for fluid channels with the porous medium as compared to the fluid channel without the porous medium. It can be concluded that the addition of the porous element in the heat sink does increase the heat transfer rate but also results in increases in the penalty of pressure drop and pumping power. Moreover, the pressure drops increases with the increase in the value of porosity. Hence the use of the porous medium to increase the heat transfer effect must be done judiciously keeping in mind the pressure drops effects.

5. Conclusion

The present study examined the performance of copper substrate inclined microjet array impingement heat sink with the addition of porous medium in a fluid channel. By varying the

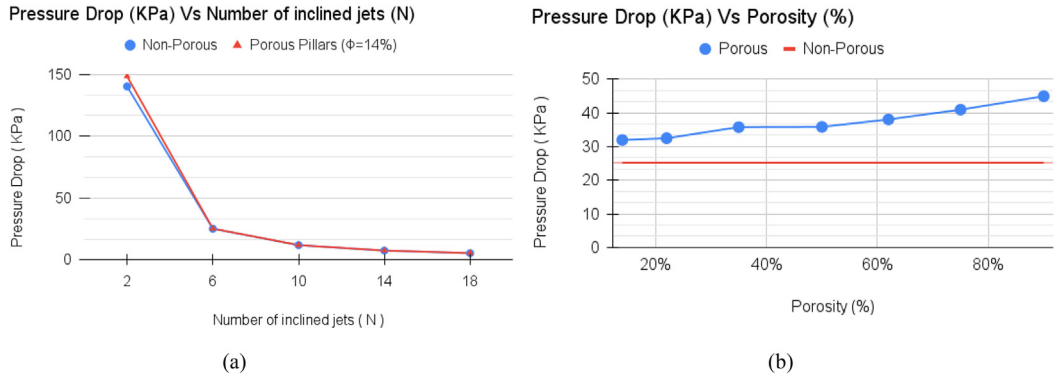


Fig. 4. Variation of pressure with (a) increasing the number of inclined jets (b) varying the porosity of the porous medium.

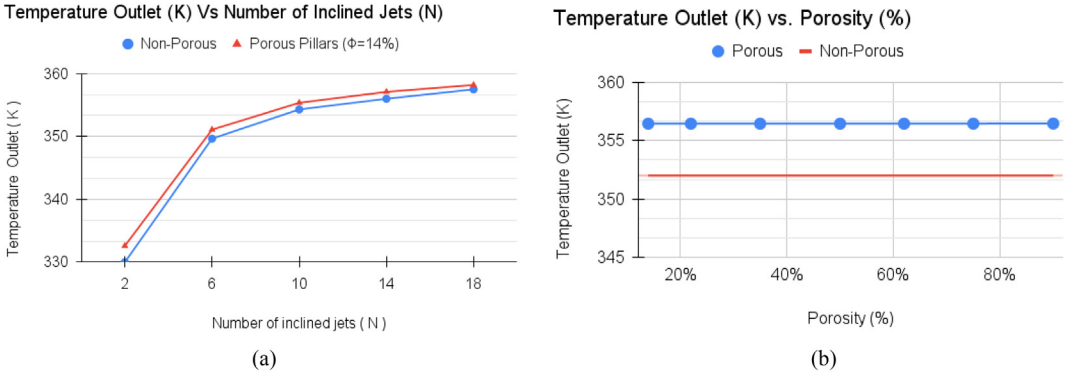


Fig. 5. Variation of Temperature Outlet with (a) increasing the number of inclined jets (b) varying the porosity of porous medium.

design parameters, such as the number of jets and the porosity of the porous medium, performance parameters such as water outlet temperature and pressure drop fluctuations were investigated. The following are the study's primary conclusions:

1. For a certain mass flow rate, the pressure drop in the inclined microjet liquid cooling system reduces as the number of jets increases. This is because as the number of jets increases, the flow rate per jet reduces, resulting in a decrease in the pressure drop value.
2. With an increase in the number of jet impinging nozzles, the output temperature rises. This indicates that the rate of heat transmission rises as the number of nozzles increases. Also, the overall temperature of the surface decreases and becomes more uniform.
3. The thermal performance of the microjet heat sink increased dramatically when porous pillars were added to the liquid channel, compared to the non-porous microjet heat sink. Although the change in the value of porosity of the porous medium did not alter the temperature significantly.
4. With the addition of porous medium and an increase in the value of porosity, the pressure drop increased. As a result, adding porous media to a microchannel should be done with caution since it improves the thermal performance at the expense of increased pumping power of heat sink.

CRediT authorship contribution statement

Harshit Jain: Conceptualization, Methodology, Software, Investigation, Validation. **Kanishk Yadav:** Visualization, Software, Writing – original draft, Investigation, Validation. **Kapeesh Kumar:** Visualization, Investigation, Software, Validation. **M. Zunaid:** Supervision, Conceptualization.

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 are grateful for the cooperation of Delhi Technological University in Delhi, India, in carrying out this research.

References

[1] D.B. Tuckerman, R.F.W. Pease, High-performance heat sinking for VLSI, IEEE Electron Dev. Lett. 2 (5) (1981) 126–129, https://doi.org/10.1109/edl.1981.25367, Institute of Electrical and Electronics Engineers (IEEE).
[2] X.F. Peng, G.P. Peterson, Convective heat transfer and flow friction for water flow in microchannel structures, Int. J. Heat Mass Transf., Elsevier BV 39 (12) (1996) 2599–2608, https://doi.org/10.1016/0017-9310(95)00327-4.
[3] T.M. Harms, M.J. Kazmierczak, F.M. Gerner, Developing convective heat transfer in deep rectangular microchannels, Int. J. Heat Fluid Flow., Elsevier BV 20 (2) (1999) 149–157, https://doi.org/10.1016/s0142-727x(98)10055-3.
[4] C.P. Tso, S.P. Mahulikar, The use of the Brinkman number for single phase forced convective heat transfer in microchannels, Int. J. Heat Mass Transf. Elsevier BV 41 (12) (1998) 1759–1769, https://doi.org/10.1016/s0017-9310(97)00232-9.
[5] T.M. Adams, S.I. Abdel-Khalik, S.M. Jeter, Z.H. Qureshi, An experimental investigation of single-phase forced convection in microchannels, Int. J. Heat Mass Transf. Elsevier BV 41 (6–7) (1998) 851–857, https://doi.org/10.1016/s0017-9310(97)00180-4.
[6] C. Yang, D. Li, J.H. Masliyah, Modeling forced liquid convection in rectangular microchannels with electrokinetic effects, Int. J. Heat Mass Transf. Elsevier BV 41 (24) (1998) 4229–4249, https://doi.org/10.1016/s0017-9310(98)00125-2.
[7] D.-Y. Lee, K. Vafai, Comparative analysis of jet impingement and microchannel cooling for high heat flux applications, Int. J. Heat Mass Transf. Elsevier BV 42 (9) (1999) 1555–1568, https://doi.org/10.1016/s0017-9310(98)00265-8.
[8] A.G. Fedorov, R. Viskanta, Three-dimensional conjugate heat transfer in the microchannel heat sink for electronic packaging, Int. J. Heat Mass Transf. Elsevier BV 43 (3) (2000) 399–415, https://doi.org/10.1016/s0017-9310(99)00151-9.
[9] X.F. Peng, B.X. Wang, G.P. Peterson, H.B. Ma, Experimental investigation of heat transfer in flat plates with rectangular microchannels, Int. J. Heat Mass Transf. Elsevier BV 38 (1) (1995) 127–137, https://doi.org/10.1016/0017-9310(94)00136-j.
[10] X.F. Peng, G.P. Peterson, The effect of thermofluid and geometrical parameters on convection of liquids through rectangular microchannels, Int. J. Heat Mass Transf. Elsevier BV 38 (4) (1995) 755–758, https://doi.org/10.1016/0017-9310(95)93010-f.
[11] X.F. Peng, G.P. Peterson, Forced convection heat transfer of single-phase binary mixtures through microchannels, Exp. Therm Fluid Sci. Elsevier BV 12 (1) (1996) 98–104, https://doi.org/10.1016/0894-1777(95)00079-8.
[12] W. Qu, Gh.M. Mala, D. Li, Heat transfer for water flow in trapezoidal silicon microchannels, Int. J. Heat Mass Transf. Elsevier BV 43 (21) (2000) 3925–3936, https://doi.org/10.1016/s0017-9310(00)00045-4.
[13] M.K. Sung, I. Mudawar, Correlation of critical heat flux in hybrid jet impingement/micro-channel cooling scheme, Int. J. Heat Mass Transf. Elsevier BV 49 (15–16) (2006) 2663–2672, https://doi.org/10.1016/j.ijheatmasstransfer.2006.01.008.
[14] J.-H. Shin, T. Rozenfeld, T. Shockner, A.K. Vutha, Y. Wang, G. Ziskind, Y. Peles, Local heat transfer under an array of micro jet impingement using HFE-7000, Appl. Therm. Eng. Elsevier BV 158 (2019) 113716, https://doi.org/10.1016/j.applthermaleng.2019.113716.
[15] X. Lu, W. Li, X. Li, J. Ren, H. Jiang, P. Ligrani, Flow and heat transfer characteristics of micro pin-fins under jet impingement arrays, Int. J. Heat Mass Transf. Elsevier BV 143 (2019) 118416, https://doi.org/10.1016/j.ijheatmasstransfer.2019.07.066.
[16] V.S. Devahdhanush, I. Mudawar, Critical Heat Flux of Confined Round Single Jet and Jet Array Impingement Boiling, Int. J. Heat Mass Transf. Elsevier BV 169 (2021) 120857, https://doi.org/10.1016/j.ijheatmasstransfer.2020.120857.
[17] C.H. Hoang, S. Rangarajan, S. Khalili, B. Ramakrisnan, V. Radmard, Y. Hadad, S. Schiffres, B. Sammakia, Hybrid microchannel/multi-jet two-phase heat sink: A benchmark and geometry optimization study of commercial product., Int. J. Heat Mass Transf. Elsevier BV 169 (2021) 120920, https://doi.org/10.1016/j.ijheatmasstransfer.2021.120920.
[18] Y. Zhang, B. Liu, J. Wei, B. Sundén, Z. Wu, Heat transfer correlations for jet impingement boiling over micro-pin-finned surface, Int. J. Heat Mass Transf. Elsevier BV 126 (2018) 401–413, https://doi.org/10.1016/j.ijheatmasstransfer.2018.04.167.
[19] X. Ju, C. Xu, Y. Zhou, Z. Liao, Y. Yang, Numerical investigation of a novel manifold micro-pin-fin heat sink combining chessboard nozzle-jet concept for ultra-high heat flux removal, Int. J. Heat Mass Transf. Elsevier BV 126 (2018) 1206–1218, https://doi.org/10.1016/j.ijheatmasstransfer.2018.06.059.
[20] S.H. Han, H.J. Park, Y.H. Kim, D.H. Lee, The effects of thermal boundary conditions on the heat transfer characteristics of laminar flow in milli-scale confined impinging slot jets, Int. J. Heat Mass Transf. Elsevier BV 168 (2021) 120865, https://doi.org/10.1016/j.ijheatmasstransfer.2020.120865.
[21] D.J. Butterfield, B.D. Iverson, D. Maynes, J. Crockett, Transient heat transfer of impinging jets on superheated wetting and non-wetting surfaces, Int. J. Heat Mass Transf. Elsevier BV 175 (2021) 121056, https://doi.org/10.1016/j.ijheatmasstransfer.2021.121056F.
[22] Y. Cui, J. Xu, K. Zhan, F. Cui, F. Hong, Bubble dynamics and pressure oscillation in highly subcooled water jet array impingement boiling under periodical heat flux, Int. Commun. Heat Mass Transf. Elsevier BV 126 (2021) 105476, https://doi.org/10.1016/j.icheatmasstransfer.2021.105476.
[23] R.D. Sundaram, S. Madhavan, P. Singh, S.V. Ekkad, Enhanced fin-effectiveness of micro-scale concentric-shape roughened target surface subjected to array jet impingement, Int. J. Heat Mass Transf. Elsevier BV 173 (2021) 121148, https://doi.org/10.1016/j.ijheatmasstransfer.2021.121148.
[24] R. Jenkins, R. Lupoi, R. Kempers, A.J. Robinson, Heat transfer performance of boiling jet array impingement on micro-grooved surfaces, Exp. Therm Fluid Sci. Elsevier BV 80 (2017) 293–304, https://doi.org/10.1016/j.expthermflusci.2016.08.006.
[25] A. Gorodetsky, T. Rozenfeld, H.D. Haustein, G. Ziskind, Flow and heat transfer analysis of hybrid cooling schemes: Adding micro-jets to a micro-gap, Int. J. Therm. Sci. Elsevier BV 138 (2019) 367–383, https://doi.org/10.1016/j.ijthermalsci.2019.01.015.
[26] A.J. Robinson, R. Kempers, J. Colenbrander, N. Bushnell, R. Chen, A single phase hybrid micro heat sink using impinging micro-jet arrays and microchannels, Appl. Therm. Eng. Elsevier BV 136 (2018) 408–418, https://doi.org/10.1016/j.applthermaleng.2018.02.058.
[27] M.V.P. Carneiro, J.R. Barbosa Jr., A comparison of parallel and colliding jet arrays in a compact vapour compression heat sink for electronics cooling, Appl. Therm. Eng. Elsevier BV 195 (2021) 117217, https://doi.org/10.1016/j.applthermaleng.2021.117217.
[28] S. Wiriyasart, P. Naphon, Heat spreading of liquid jet impingement cooling of cold plate heat sink with different fin shapes, Case Stud. Therm. Eng. Elsevier BV 20 (2020), https://doi.org/10.1016/j.csite.2020.100638.
[29] P. Naphon, L. Nakharintr, S. Wiriyasart, Continuous nanofluids jet impingement heat transfer and flow in a micro-channel heat sink, Int. J. Heat Mass Transf. Elsevier BV 126 (2018) 924–932, https://doi.org/10.1016/j.ijheatmasstransfer.2018.05.101.

[30] P. Singh, M. Zhang, S. Ahmed, K.R. Ramakrishnan, S. Ekkad, Effect of micro-roughness shapes on jet impingement heat transfer and fin-effectiveness, Int. J. Heat Mass Transf. Elsevier BV 132 (2019) 80–95, https://doi.org/10.1016/j.ijheatmasstransfer.2018.11.135.
[31] L. Gong, J. Zhao, S. Huang, Numerical study on layout of micro-channel heat sink for thermal management of electronic devices, Appl. Therm. Eng. Elsevier BV 88 (2015) 480–490, https://doi.org/10.1016/j.applthermaleng.2014.09.048.
[32] H.C. Cui, X.T. Lai, J.F. Wu, M.Z. Wang, W. Liu, Z.C. Liu, Overall numerical simulation and experimental study of a hybrid oblique-rib and submerged jet impingement/microchannel heat sink, Int. J. Heat Mass Transf. Elsevier BV 167 (2021), https://doi.org/10.1016/j.ijheatmasstransfer.2020.120839.
[33] D. Lelea, The heat transfer and fluid flow of a partially heated microchannel heat sink, Int. Commun. Heat Mass Transf. Elsevier BV 36 (8) (2009) 794–798, https://doi.org/10.1016/j.icheatmasstransfer.2009.05.003.
[34] P. Naphon, S. Wongwises, Investigation on the jet liquid impingement heat transfer for the central processing unit of personal computers, Int. Commun. Heat Mass Transf. Elsevier BV 37 (7) (2010) 822–826, https://doi.org/10.1016/j.icheatmasstransfer.2010.05.004.
[35] M. Kumar, J. Rajiyan, P. Gupta, A computational approach for solving elasto-statics problems, Mater. Today:. Proc. Elsevier BV 46 (2021) 6876–6879, https://doi.org/10.1016/j.matpr.2021.04.462.
[36] Arun, K. Karunanithi., Fatemeh Hassanipour, "Heat Sink With Stacked Multi-Layer Porous Media," Proceedings of the 5th International Conference on Porous Media and its Applications in Science and Engineering ICPM5 june 22–27 (2014), Kona, Hawaii, USA.
[37] R.A. Mahdi, H.A. Mohammed, K.M. Munisamy, N.H. Saeid, Influence of various geometrical shapes on mixed convection through an opencell aluminium foam filled with nano fluid, J. Comput. Theor. Nanosci. 11 (5) (2014) 1275–1289.
[38] M. Zunaid, A. Husain, B. Singh Chauhan, R. Sahu, Numerical analysis of inclined jet impingement heat transfer in microchannel, Mater. Today Proc. Elsevier BV 43 (2021) 557–563, https://doi.org/10.1016/j.matpr.2020.12.048.

Performance Evaluation of Three-Phase Unified Power Quality Conditioner controlled using SRF

Rahul Singh
Dept. of Electrical Engg.
Delhi Technological University
New Delhi, India
rahul164singh@gmail.com

Alka Singh
Dept. of Electrical Engg.
Delhi Technological University
New Delhi, India
alkasingh.eed@gmail.com

2021 4th International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE) | 978-1-6654-1429-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/RDCAPE52977.2021.9633640

***Abstract*—This paper discusses the performance evaluation of three-phase three-wire Unified Power Quality Conditioner (UPQC) using Synchronous Reference Frame (SRF) Theory. The core objective of any electrical utility is to supply quality power to load end uninterruptedly with constant frequency, sinusoidal balanced rms voltage and current. Electrical and electronic equipment such as variable speed drives, Switch-mode power supplies (SMPS), high penetration of non-linear loads injects lot of harmonics in supply current of the system thereby affects the neighboring loads and source as well. On the other hand, compressors, induction motors, pumps etc. draw high input surge current during starting thereby creating sag/swell in system voltage. To compensate power quality concerns associated with both current and voltage, a device called UPQC considered to be the most prevalent device for sensitive and critical loads. It consists of shunt and series Active Power Filters (APF) which are connected back-to-back using dc-link capacitor. A PI controller is used for the balancing of dc-link capacitor voltage. UPQC is used to alleviate harmonics in load voltage, supply current; and balances the load voltage of the distribution system. In this paper, UPQC is designed, modeled, controlled and simulated for different disturbances in source voltage i.e. sag/swell, polluted grid. Results are presented to analyze the total harmonic distortion (THD) in voltage and current, power flow and dc-link voltage for nonlinear load using Simulink based MATLAB model.**

Keywords—APF, Phase-locked loop (PLL), SRF, THD, UPQC, Point of common coupling (PCC)

I. INTRODUCTION

An electrical utility must provide quality power to its customers and failure to do so might result in equipment malfunction, permanent breakdown or poor performance of the equipment. With rapid increase in nonlinear power electronics loads, load current has high harmonic content. Large harmonic content may lead to failure of sophisticated electronic equipment, damage to capacitors, heating of electrical equipment, resonance phenomena, flickering of lights etc. While the unbalance/ distortion in voltage at load side may lead to increase in burden on distribution transformer during voltage swell condition, overheating and insulation breakdown of transformer windings, failure of electrical and electronic equipment. As per IEEE 1159-2019 [1] "a typical duration of instantaneous swell/sag is 0.5 to 30 cycles". Heavy loads such as motor drives, compressors etc. draw high surge current which may cause sags in voltage and lead to failure of transmission and distribution system, It is also observed that in most of the cases sudden rise in loading, short circuit faults leads to sag in voltage whereas sudden load shedding leads to swell in voltage. Non-linear loads will introduce harmonic currents into the grid owing to system impedance, harmonic currents will distort the voltage waveform at the PCC [2]. These power quality issues may lead to complete failure of the grid. Such voltage quality issues frequently result in multiple erroneous tripping of power electronic systems, failure and, excessive capacitor bank heating [3]. To eliminate such power quality issues UPQC may be installed at PCC in transmission and distribution system. UPQC comprises a back-to-back connection of series and shunt APF with a dc-link capacitor [4] functioning as an interface between them to efficiently control both current and voltage related power quality concerns [5]. The major advantage of UPQC is that, as per IEEE 519-2014 standard [6] it can effectively reduce the THD less than 5 %in the voltage and current even when the utility grid is polluted.

In the subsequent sections of this paper configuration & designing of UPQC are discussed using SRF theory [7], [8] and, dynamic performance is evaluated for different disturbances along with the analysis of capacitor dc-link voltage, power flow when system is connected to nonlinear load. THD in current and voltage is analyzed for polluted grid. Conclusions based on the MATLAB-Simulink results are provided in the last section.

II. SYSTEM CONFIGURATION

UPQC is a device used in distribution systems to enhance power quality. It is comprised of a shunt and a series APF, are linked together in a cascaded manner using a dc-link capacitor. At the PCC, a series transformer of 1:1 turns ratio is used to connect the series APF and, a coupling inductor is used to connect shunt APF. In this configuration series and shunt APF is connected towards the source and load side respectively. By using ripple filters the induced harmonics due to the converters switching operation are filtered out. For the operation of both the APFs simultaneously, two reference signals are generated viz. for source current and load voltage. SRF theory is employed to control UPQC. Hysteresis Controller (HC) is used to generate the switching pulses to operate the voltage source converter (VSC) with insulated gate bipolar transistor (IGBT) [9] to inject compensating voltage and current into the system. The HC generates switching pulses quickly, efficiently and with minimal complexity [10].

III. CONTROL SCHEME OF UPQC

This section addresses the use of SRF theory in designing of UPQC series and shunt APF controllers. SRF theory deals with the transformation of abc into the dq0 frame, i.e. Park's transformation, to obtain the reference load voltage and source current signals.

A. Control Scheme of Series APF

In this control scheme, the sensed source voltage V_{sabc} is transformed into dq coordinates using Park's transformation as shown in (1).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:02:44 UTC from IEEE Xplore. Restrictions apply.

$$\begin{bmatrix} V_{sd} \\ V_{sq} \end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix} \sin(\theta) & \sin(\theta - \frac{2\pi}{3}) & \sin(\theta + \frac{2\pi}{3}) \\ \cos(\theta) & \cos(\theta - \frac{2\pi}{3}) & \cos(\theta + \frac{2\pi}{3}) \end{bmatrix} \begin{bmatrix} V_{sa} \\ V_{sb} \\ V_{sc} \end{bmatrix} \quad (1)$$

Fig. 1 illustrates the control schematic for series APF using SRF theory. The transformed voltage V_{sd} and V_{sq} contains dc as well as ac component. LPF is utilised to get the dc fundamental components, i.e. V_{qdc} and V_{ddc}. For the synchronization PLL is used. To regulate the ac magnitude of load voltage V_l to its reference voltage V_l^*, a PI controller is employed.

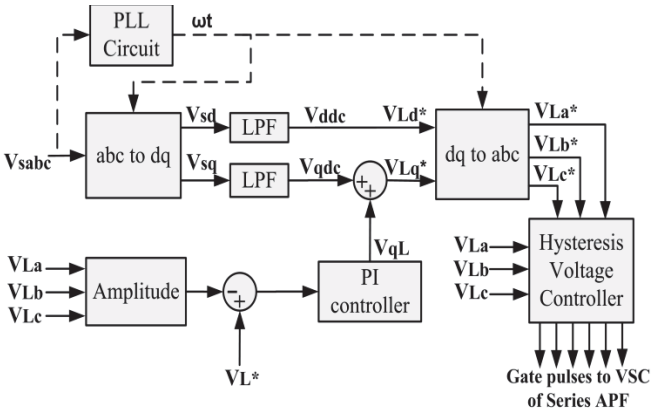


Fig. 1. Control schematic of Series APF

Where, load phase voltages V_{la} , V_{lb} , V_{lc} are used to compute the amplitude of V_l given by (2) and reference voltage V_l^* is given in (3).

$$V_l = \sqrt{\frac{2}{3}(V_{la}^2 + V_{lb}^2 + V_{lc}^2)} \quad (2)$$

$$V_l^* = \sqrt{\frac{2}{3}} \times (\text{desired rms value of load voltage}) \quad (3)$$

To get q-axis reference load voltage V_{Lq}^* the PI controller's controlled signal V_{ql} is added to the V_{qdc} signal given by (4), d-axis reference load voltage V_{Ld}^* is given in (5)

$$V_{Lq}^* = V_{qdc} + V_{ql} \quad (4)$$

$$V_{Ld}^* = V_{ddc} \quad (5)$$

On applying the inverse Park's transformation using (6) gives the reference load voltage i.e. V_{la}^*, V_{lb}^*, V_{lc}^*.

$$\begin{bmatrix} V_{la}^* \\ V_{lb}^* \\ V_{lc}^* \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \cos(\theta - \frac{2\pi}{3}) & -\sin(\theta - \frac{2\pi}{3}) \\ \cos(\theta + \frac{2\pi}{3}) & -\sin(\theta + \frac{2\pi}{3}) \end{bmatrix} \begin{bmatrix} V_{Ld}^* \\ V_{Lq}^* \end{bmatrix} \quad (6)$$

For generating the triggering pulses to derive the converter of series APF The sensed load voltage V_{labc} is compared using hysteresis voltage controller with reference load voltage V_{labc}^*.

B. Control Scheme of Shunt APF

In this control scheme three-phase load current i_{labc} is sensed, transformed from abc coordinates to dq frame is shown in (7). For the implementation of the SRF method, synchronizing system i.e. phase-locked loop (PLL) should be used. The transformed load current i_{ld} , i_{lq} comprises two components the direct and quadrature axis components. d-axis current is the combination of DC part consists the fundamental component (i_{ddc}) and AC part consists of harmonic component of current (i_h) shown in (8). Further to remove the harmonics present in the transformed dq signal, "low pass filter" (LPF) of 2nd order with corner frequency of 20Hz are used [11]. The output from the filters is only the fundamental component and free from lower order harmonics. Fig. 2 illustrates the control schematic of shunt APF.

$$\begin{bmatrix} i_{ld} \\ i_{lq} \end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix} \sin(\theta) & \sin(\theta - \frac{2\pi}{3}) & \sin(\theta + \frac{2\pi}{3}) \\ \cos(\theta) & \cos(\theta - \frac{2\pi}{3}) & \cos(\theta + \frac{2\pi}{3}) \end{bmatrix} \begin{bmatrix} i_{la} \\ i_{lb} \\ i_{lc} \end{bmatrix} \quad (7)$$

$$i_{ld} = i_{ddc} + i_h \quad (8)$$

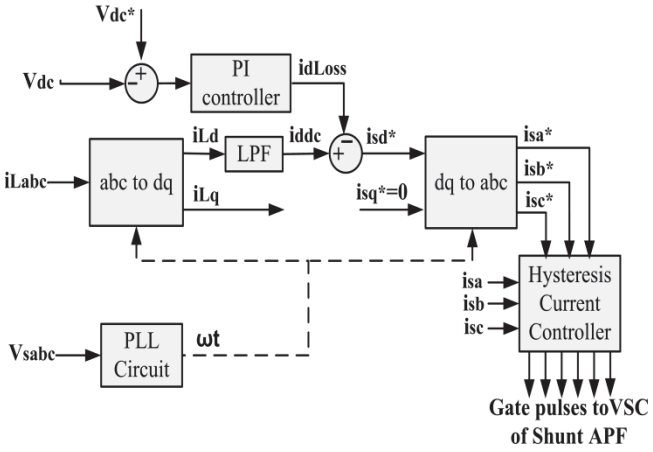


Fig. 2. Control schematic of Shunt APF

Shunt APF has one of the primary function is to maintain a stable dc-link voltage V_{dc} to preserve equilibrium of power in system. Using a suitable PI controller "V_{dc}" is kept constant. It also determines the component of loss current (i_{dloss}) caused by power imbalance in the system shown in (9).

$$i_{sd}^* = i_{ddc} - i_{dloss} \quad (9)$$

$$\begin{bmatrix} i_{sa}^* \\ i_{sb}^* \\ i_{sc}^* \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \cos(\theta - \frac{2\pi}{3}) & -\sin(\theta - \frac{2\pi}{3}) \\ \cos(\theta + \frac{2\pi}{3}) & -\sin(\theta + \frac{2\pi}{3}) \end{bmatrix} \begin{bmatrix} i_{sd}^* \\ i_{sq}^* \end{bmatrix} \quad (10)$$

Hence, by applying the inverse Park's transformation using (10) gives the reference source current i.e. i_{sa}^*, i_{sb}^*, i_{sc}^*. The sensed source current i_{sa} , i_{sb} , i_{sc} are compared using hysteresis current controller with reference source current to provide triggering pulses to derive the converter of shunt APF.

IV. Simulation Results

The Matlab based simulation model for UPQC connected to the distribution system to enhance the power quality during sag/swell, polluted voltage and system containing current harmonics is shown in Fig. 3. In this section UPQC dynamic performance is evaluated for different disturbances along with the analysis of capacitor dc-link voltage, power flow when

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:02:44 UTC from IEEE Xplore. Restrictions apply.

system is connected to nonlinear load. THD in source current and load voltage are analyzed for polluted grid. These cases are discussed as follows.

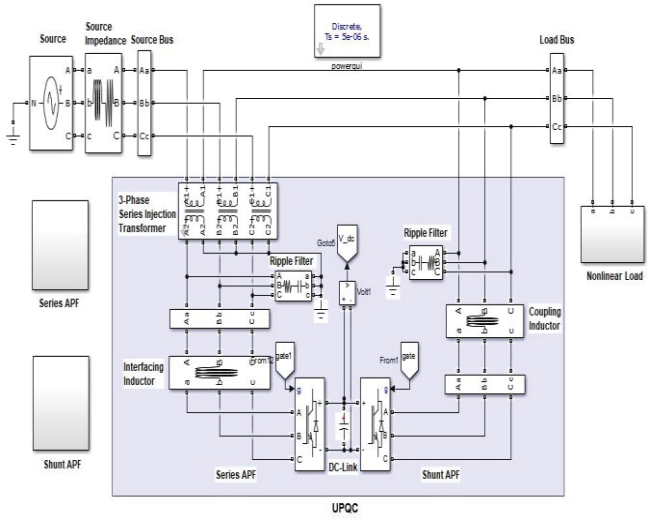


Fig. 3. Matlab based Simulation model for UPQC

A. Analysis during Voltage Sag under Ideal Grid

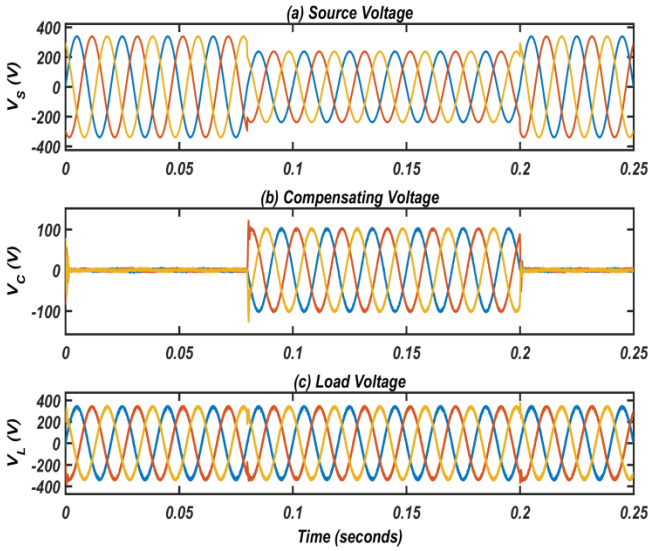


Fig. 4. (a) source (b) compensating (c) load voltage during sag

Here in source voltage, 30% sag is given for the time interval of 0.08s to 0.2s depicted by Fig. 4(a). Series APF of UPQC injects compensating voltage during sag interval depicted by Fig. 4(b) to maintain sinusoidal load voltage shown in Fig. 4(c). A fall in capacitor dc-link voltage is observed due to sudden sag introduction. The shunt APF will deliver some active power to series APF via dc-link capacitor by absorbing it from the source bus. This leads to absorb more current from source terminals for power balance of the system. Therefore, this causes an increase in compensated source current, shown in Fig. 5(a) during the sag interval. Shown in Fig. 5(b) load current which is clearly distorted. Depicted in Fig. 5(c) load harmonic current demand, is provided by shunt APF. Both the VSC's are functioning well, the source current is sinusoidal and, the load voltage is well regulated. Fig. 6 presents the system power flow during voltage sag.

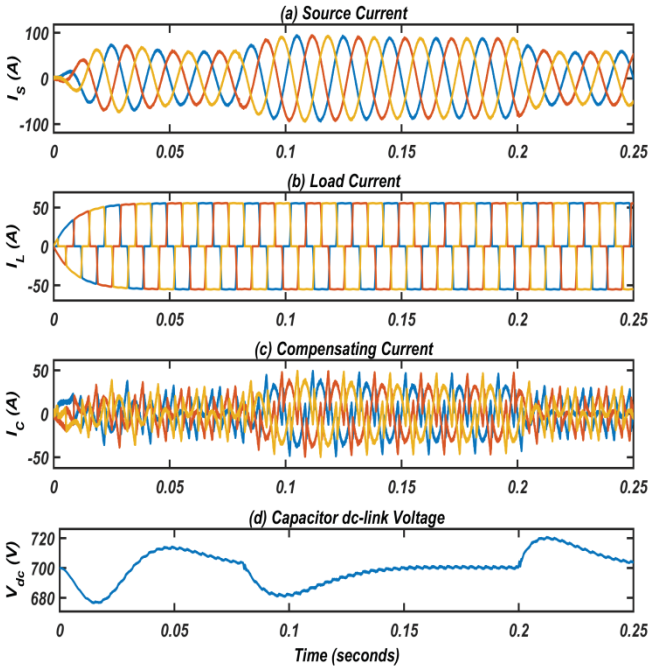


Fig. 5. (a) source current (b) load current (c) compensating current and (d) capacitor dc-link voltage during sag

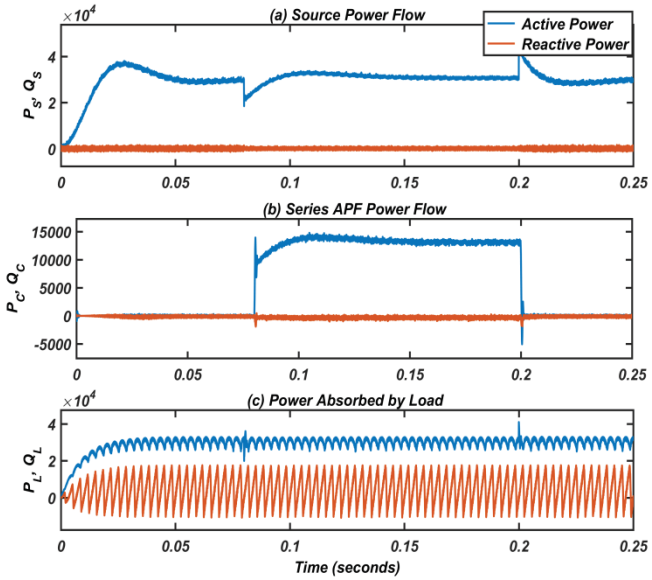


Fig. 6. Active & Reactive power flow of (a) source (b) Series APF (c) load, during sag in voltage

B. Analysis during Voltage Swell under Ideal Grid

Here in source voltage, 30% swell is given for the time interval of 0.08s to 0.2s depicted by Fig. 7(a). During swell, as presented in Fig. 7(b) series APF absorbs excess voltage to maintain balanced sinusoidal load voltage as presented in Fig. 7(c). During voltage swell condition, a rise in capacitor dc-link voltage detected. Thereby, it will deliver some amount of active power to shunt APF by absorbing it from the source bus via capacitor dc-link which is used for lowering the harmonics in source current by injecting harmonic current demand of the load. This leads to fall in compensated source current as shown in Fig. 8(a) for swell interval. This is vital to maintain the system's power balance. Fig. 8(b) shows the load current which is distorted. Simultaneously, shunt APF provides the load harmonic current demand, represented by

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:02:44 UTC from IEEE Xplore. Restrictions apply.

Fig. 8(c). DC-link voltage is regulated by the PI controller to a predetermined reference value of 700V.

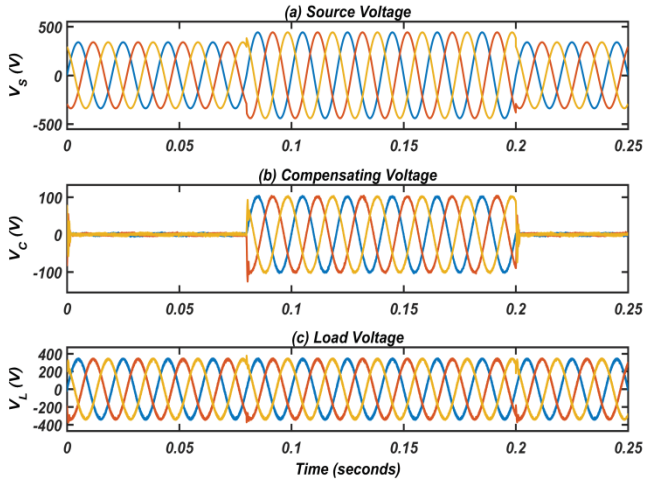


Fig. 7. (a) source (b) compensating (c) load voltage during swell

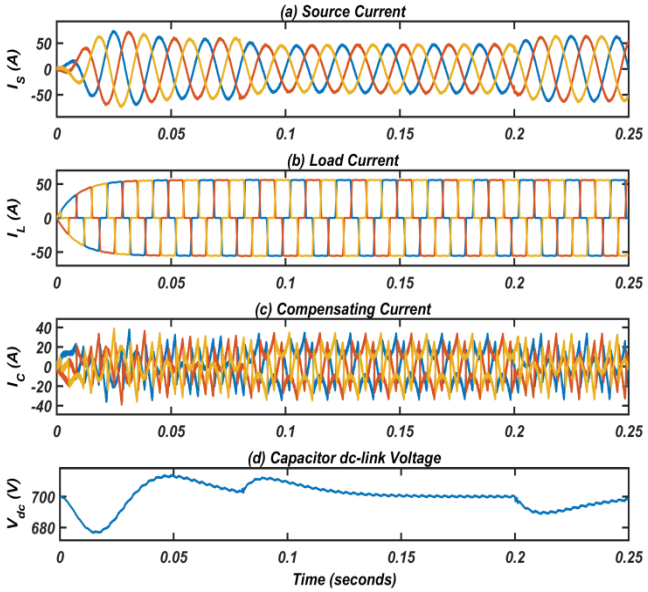


Fig. 8. (a) source current (b) load current (c) compensating current and (d) capacitor dc-link voltage during swell

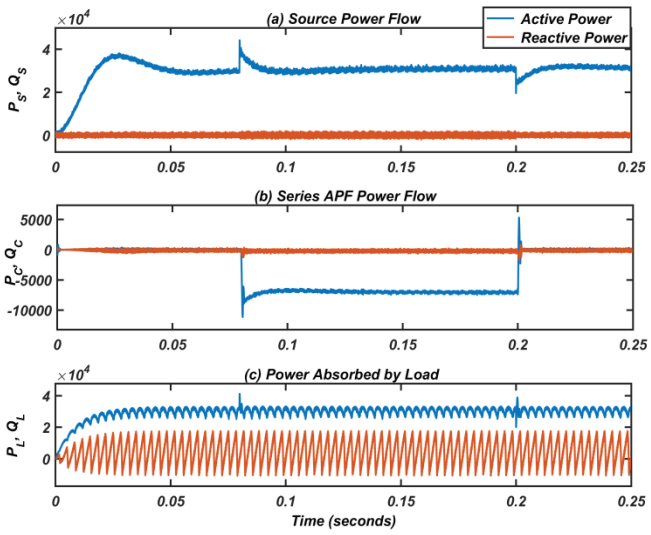


Fig. 9. Active & Reactive power flow of (a) source (b) Series APF (c) load, during swell in voltage

The Fig. 9 shows power flow of the system during voltage swell

C. *Analysis during Polluted Source Voltage*

In this case, with "5th & 7th" order harmonics of magnitude "20 % & 15 %" correspondingly, source voltage is polluted, for the time interval of 0.05s to 0.2s as shown in Fig. 10(a) i.e. with THD of 25%. When UPQC is not connected to the system, nonlinear load from the polluted supply draws non-sinusoidal current, affecting the surrounding loads and the sources too, THD in current is 26.65%. Now, when UPQC is there in the system, the Series APF will start injecting compensated voltage component shown in Fig. 10(b) into the system to alleviate the harmonics in load voltage. The series APF can lower the THD content of load voltage to 2.87%. Whereas, shunt APF will supply deviated fundamental current component shown in Fig. 11(c) to lower the source current THD to 3.33%. Fig. 11(a) and 10(c) shows the compensated source current and load voltage respectively. Capacitor dc-link voltage is represented by Fig. 11(d) that remains balanced by using the shunt APF's PI controller. Throughout the compensation process the UPQC supplies the reactive power to nonlinear load to maintain source current in-phase.

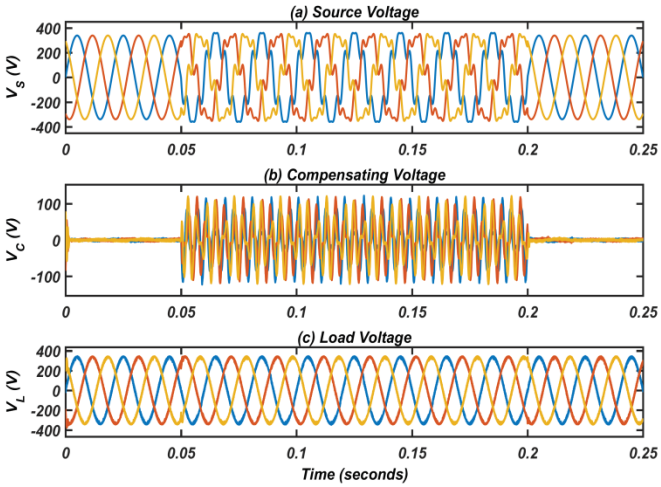


Fig. 10. (a) source (b) compensating (c) load voltage ;when source is polluted

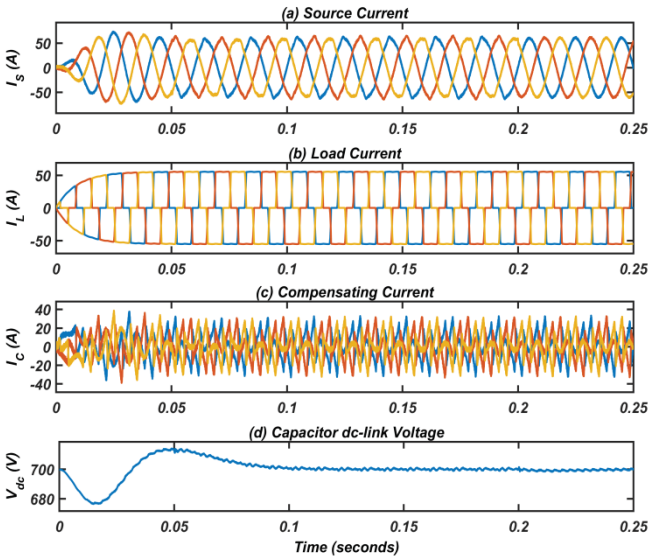


Fig. 11. (a) source current (b) load current (c) compensating current and (d) capacitor dc-link voltage, when source is polluted

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:02:44 UTC from IEEE Xplore. Restrictions apply.

Fig. 12 shows the active & reactive power flow during polluted source voltage.

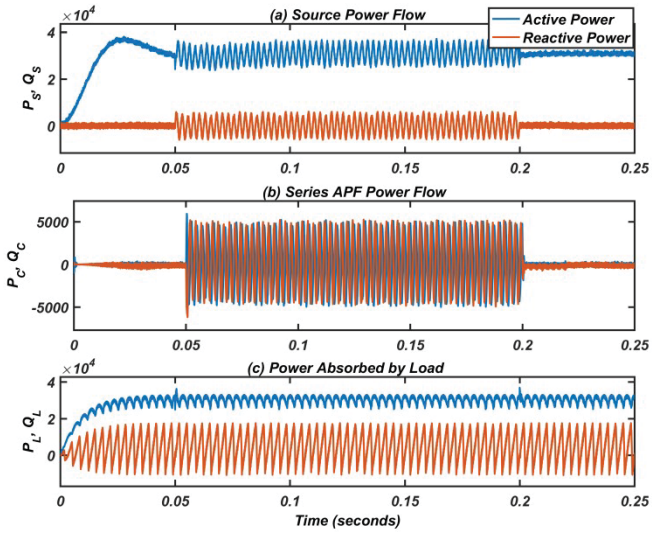


Fig. 12. Active & Reactive power flow of (a) source (b) Series APF (c) load, when source voltage is polluted

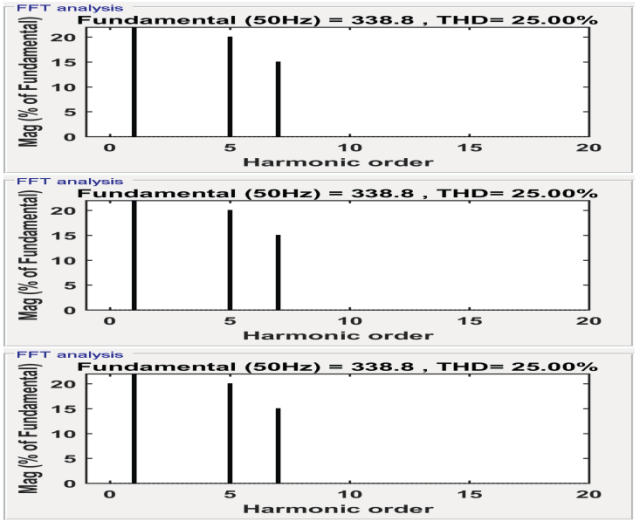


Fig. 13. THD in load phase voltages before compensation

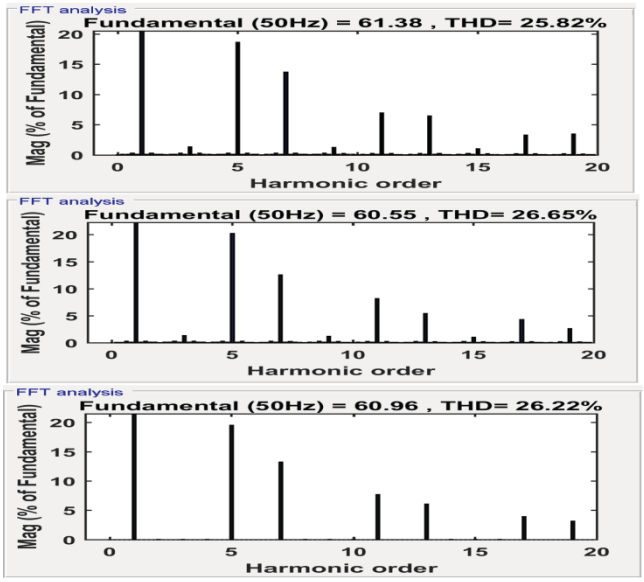


Fig. 14. THD in source phase current before compensation

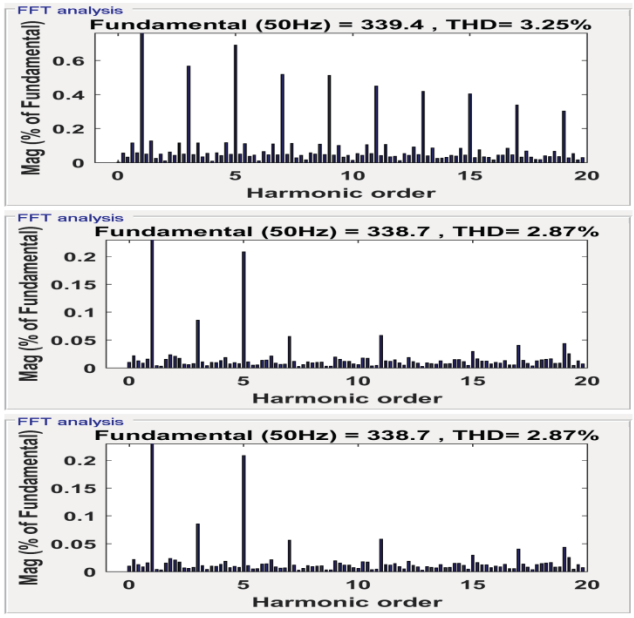


Fig. 15. THD in load phase voltages after compensation

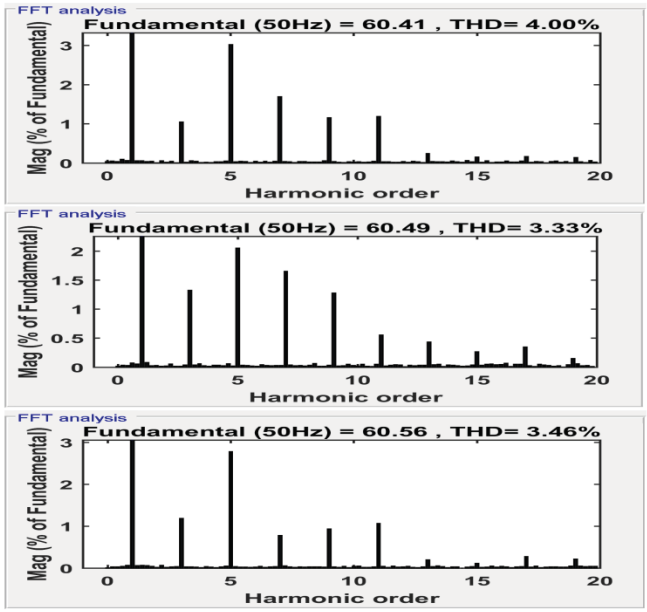


Fig. 16. THD in source phase current after compensation

Fig. 13, 14, 15 and, 16 shows the THD in load phase voltage and source phase current prior-to and post compensation which are 25%, 26.65% and 2.87%, 3.33% respectively.

TABLE. I COMPARISON OF MAGNITUDE AND THD IN LOAD VOLTAGE & SOURCE CURRENT WITHOUT & WITH UPQC

Parameters		Without UPQC		With UPQC	
		Magnitude	THD (in %)	Magnitude	THD (in %)
Load Voltage	V_{La}	338.8 V	25	339.4 V	3.25
	V_{Lb}	338.8 V	25	338.7 V	2.87
	V_{Lc}	338.8 V	25	338.7 V	2.87
Source Current	I_{sa}	61.38 A	25.82	60.41 A	4
	I_{sb}	60.55 A	26.65	60.49 A	3.33
	I_{sc}	60.96 A	26.22	60.56 A	3.46

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:02:44 UTC from IEEE Xplore. Restrictions apply.

Table 1 shows the comparison of magnitude and THD in voltage and current without and with UPQC. THD in both load voltage & source current has been minimized to a level of "<5%" with the application of UPQC.

D. *Analysis of DC-Link Voltage during Sag/Swell and for Polluted Source*

When source voltage is free from sag/swell and harmonics then, it will only supply active power whereas, UPQC supplies the reactive power and harmonic current demand of the nonlinear load. Fig. 17 depicts the capacitor dc-link voltage (V_{dc}) is maintained constant.

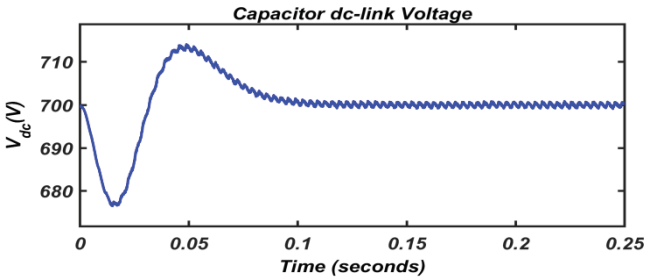


Fig. 17. Capacitor dc-link voltage under normal operation

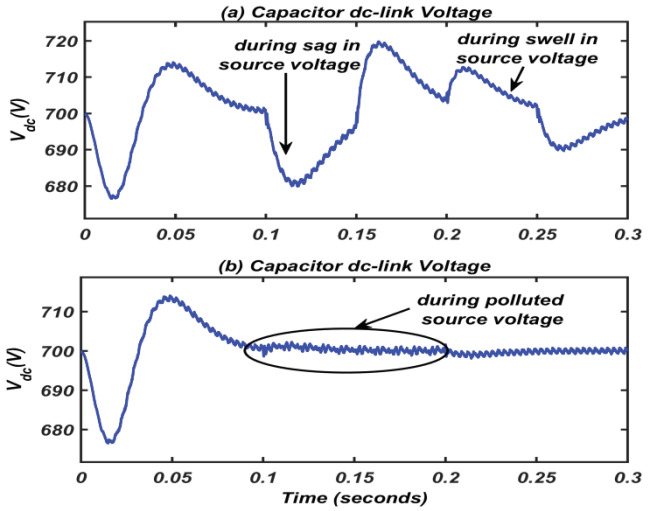


Fig. 18. (a) Capacitor dc-link voltage during; (a) sag/swell, (b) polluted source voltage

Fig. 18(a) shows when source voltage is having sag/swell then, voltage at DC link changes. It is brought under control using PI regulator. During sag condition, dc-link voltage V_{dc} falls, series APF absorbs the active power from the shunt APF via capacitor dc-link to supply nonlinear load, while shunt APF draws that active power from the supply utility. And, during swell dc-link voltage V_{dc} rises, series APF delivers the active power to shunt APF via capacitor dc-link to supply nonlinear load, while series APF draws that active power from the supply utility. The complete reactive power is supplied by the shunt APF to meet the demand of the load. Fig. 18(b) shows the trivial rise in V_{dc} for time interval when source voltage is having harmonics. UPQC supplies compensating harmonic voltage and current in the system to balance source current and load voltage.

V. Conclusion

The three-phase UPQC control scheme and dynamic performance have been investigated for nonlinear loading when supply utility is having sag/swell and distortion in voltage. From the results, it is observed that UPQC effectively compensates the sag/swell related problem to balance the load voltage and, mitigate the harmonics in voltage and current of the grid as per the prescribed IEEE standards. The PI controller of UPQC's shunt APF is well regulated, the dc-link voltage is controlled for system power balance during sag/swell, such that UPQC supplies reactive power and, the supply utility supplies only active power to the load. In this paper, three cases showing swell, sag in supply grid and polluted grid have been considered. The control strategy proposed for the shunt and series VSC works well.

Appendix

1. Supply voltage : 415V rms , 50Hz
2. Source resistance and inductance: 0.1mΩ, 1μH
3. dc-link capacitor = 5500μF, 700V
4. Shunt APF parameters: Coupling inductance=1mH, Ripple filter Rc=10Ω, Cc =10μF, K_P= 1 & K_I= 40
5. Series APF parameters: Injection series transformer Rating=14.28KVA, 1:1 turns ratio, Ripple filter Rc=1Ω, Cc=30μF, Interfacing Inductor= 8mH, K_P= 1.5& K_I= 0.8
6. Load : Nonlinear R_L=10, L=90mH

References

[1] "IEEE recommended practice for monitoring electric power quality," IEEE, Piscataway, NJ, USA, 2019.

[2] Q. N. Trinh and H. H. Lee, "Improvement of power quality under distorted source and nonlinear load conditions," in *Proceedings of The 7th International Power Electronics and Motion Control Conference*, 2012.

[3] S. Devassy and B. Singh, "Design and performance analysis of three-phase solar PV integrated UPQC," *IEEE Trans. Ind. Appl.*, vol. 54, no. 1, pp. 73–81, 2018.

[4] S. Vinnakoti and V. R. Kota, "Implementation of artificial neural network based controller for a five-level converter based UPQC," *Alex. Eng. J.*, vol. 57, no. 3, pp. 1475–1488, 2018.

[5] K. Hasan, M. M. Othman, N. F. A. Rahman, M. A. Hannan, and I. Musirin, "Significant implication of unified power quality conditioner in power quality problems mitigation," *Int. J. Power Electron. Drive Syst. (IJPEDS)*, vol. 10, no. 4, p. 2231, 2019.

[6] "IEEE recommended practice and requirements for harmonic control in electric power systems," IEEE, Piscataway, NJ, USA, 2014.

[7] Y. Pal, A. Swarup, and B. Singh, "A comparative analysis of three-phase four-wire UPQC topologies," in *2010 Joint International Conference on Power Electronics, Drives and Energy Systems & 2010 Power India*, 2010.

[8] D. Krishna, M. Sasikala and V. Ganesh, "Mathematical modeling and simulation of UPQC in distributed power systems", *Proc. - 2017 IEEE Int. Conf. Electr. Instrum. Commun. Eng. ICEICE 2017*, vol. 2017-Decem, no. 3, pp. 1-5, 2017.

[9] M. Linca, C. V. Suru, and C. A. Preda, "Indirect current control algorithm implementation for an active filtering system using constant switching frequency hysteresis controllers," in *2018 International Conference on Applied and Theoretical Electricity (ICATE)*, 2018.

[10] N. Kimkaew and P. Santiprapan, "The performance comparison of switching techniques for unified power quality conditioner," in *2021 9th International Electrical Engineering Congress (iEECON)*, 2021.

[11] N. Zaveri, A. Mehta, and A. Chudasama, "Performance analysis of various SRF methods in three phase shunt active filters," in *2009 International Conference on Industrial and Information Systems (ICIIS)*, 2009.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:02:44 UTC from IEEE Xplore. Restrictions apply.

Applied Physics A (2022) 128:28
https://doi.org/10.1007/s00339-021-05096-2

Applied Physics A
Materials Science & Processing

Plasma-based nanoarchitectonics for vertically aligned dual-metal carbon nanotube field-effect transistor (VA-DMCNFET) device: effect of plasma parameters on transistor properties

Mansha Kansal[1] · **Suresh C. Sharma**[1]

Received: 4 October 2021 / Accepted: 8 November 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH, DE part of Springer Nature 2021

Abstract
In this paper, we intend to show the role of plasma parameters on the performance of the simulated device Plasma-Assisted Vertically Aligned Dual-Metal Carbon Nanotube Field-Effect Transistor (VA-DMCNFET) for the first time. Here, vertically aligned semiconducting Carbon Nanotube (CNT) synthesized using Plasma-Enhanced Chemical Vapor Deposition (PECVD) technique is implemented as the channel. The proposed device shows significant improvement in performance over conventional Nanowire Field-Effect Transistor (NWFET). Next, the device geometry and electrical properties are related to the PECVD parameters by means of a physical model. The DC and analog performance of VA-DMCNFET at various plasma parameters corresponding to different values of CNT channel radius (for fixed CNT channel length) and different values of CNT channel length (for fixed CNT channel radius) is analyzed. It is observed that lower values of plasma parameters are essential for higher values of drain current, transconductance, output conductance, cutoff frequency, and lower values of threshold voltage and channel resistance. On the other hand, higher values of plasma parameters are essential for better I_{ON}/I_{OFF} current ratio, early voltage, and gain of the proposed device. By altering the plasma parameters, efficiency of the device can be improved leading to better real-life applicability and performance. The results obtained from this study have been verified with the existing experimental observations for validation of chosen mechanisms.

Keywords Plasma parameters · Vertically aligned carbon nanotube · Carbon nanotube field-effect transistor · Plasma-enhanced chemical vapor deposition

1 Introduction

The semiconductor industry is faced with a daunting problem that has invited extensive research so far–reduce the physical dimensions while improving the performance of the device [1]. The scaling limitations in silicon are being considered by integration of non-conventional high mobility channel materials [2, 3]. Carbon Nanotube (CNT) has been extensively researched lately with the same intention. CNTs are characterized as graphitic filaments which have a cylindrical shape of radius in the nanometer scale and length in the micrometer range [4, 5].

CNT field-effect transistors were first showcased in 1998 [6]. Since then, they have been the subject of significant research interest leading to significant progress in understanding device physics and in enhancing transistor performance due to high speed operation of CNT [7, 8]. The CNTFETs show superior device performance in comparison with existing silicon-based devices because of ballistic transport in the single-walled CNTs.

The Plasma-Enhanced Chemical Vapor Deposition (PECVD) technique gives birth to vertically aligned carbon nanostructures. Growth of vertically aligned CNT is due to the electric field of the plasma sheath which applies force on the nanostructure, leading to growth in a vertical direction [9–14]. PECVD technique stands out from the rest because the resultant CNTs produced have a high semiconductive CNT ratio leading to higher values of drain current [15]. Since CNTFETs are involved in a wide range of

✉ Suresh C. Sharma
prof_sureshsharma@dtu.ac.in

Mansha Kansal
kansal.mansha@gmail.com

[1] Department of Applied Physics, Delhi Technological University, Bawana Road, Delhi 110085, India

Published online: 10 December 2021

applications, the said ratio is of paramount importance. For CNTs synthesized using PECVD, this can be as high as 90% [16, 17]. Plasma-based nanoarchitectonics deal with the production of systems from plasma-based nanoscale objects. Nanoarchitectonics will yield application opportunities ensuring these nanomaterials benefit our day-to-day lives.

Mizutani et al. [18] have fabricated CNTFETs using the grid inserted PECVD technique. The horizontally aligned nanotubes were used as a channel. With a channel length of 7 µm and channel width of 30 µm, they were able to obtain an ON/OFF current ratio of 10^3 and good pinch-off characteristics. Bashir et al. [19] have proposed a novel device of Schottky barrier metal–oxide–semiconductor field-effect transistor (MOSFET), where Hafnium creates charge plasma in the undoped silicon. Due to the employment of charge plasma concept, various properties like I_{ON}/I_{OFF} current ratio, ON current, cutoff frequency have significantly improved.

In this paper, for the first time, the role of plasma parameters on the performance of the simulated Plasma-Assisted Vertically Aligned Dual-Metal Carbon Nanotube Field-Effect Transistor (VA-DMCNFET) with cylindrical dual-metal surrounding gate is studied. From the obtained results, it can be observed that VA-DMCNFET shows better performance over Nanowire Field-Effect Transistor (NWFET). It is further shown that lower values of plasma parameters are essential for higher values of drain current, transconductance, output conductance and cutoff frequency, and lower values of threshold voltage and channel resistance. On the other hand, higher values of plasma parameters are essential for better values of I_{ON}/I_{OFF} current ratio, early voltage and gain of the proposed device. Hence, finding the right balance between the appropriate plasma parameters becomes important for optimum device performance.

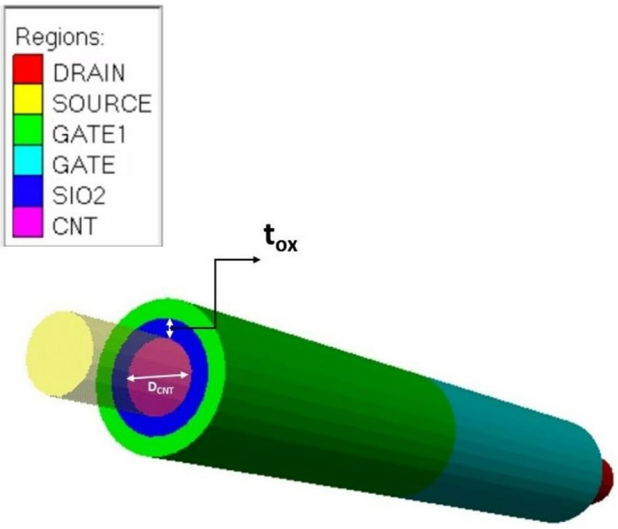


Fig. 1 3-D structure of cylindrical VA-DMCNFET

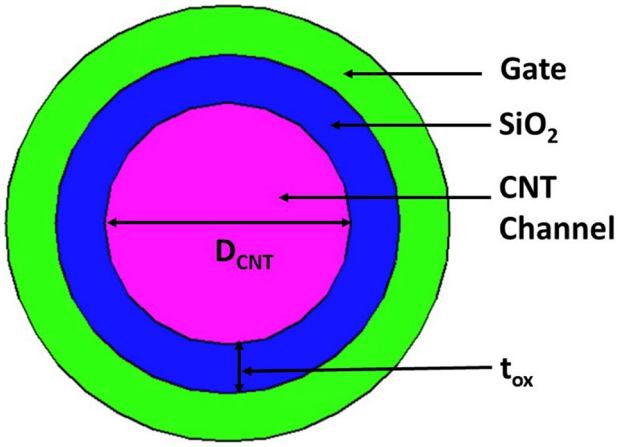


Fig. 2 Cross-sectional view of VA-DMCNFET

2 Device simulation model

The proposed device Plasma-Assisted Vertically Aligned Dual-Metal Carbon Nanotube Field-Effect Transistor (VA-DMCNFET) is simulated. The carbon nanotubes have been experimentally synthesized via microwave PECVD using a mixture of argon and methane gases on nickel coated and silicon substrates [20]. Figure 1 and Fig. 2 represent the 3-D and cross-sectional view of VA-DMCNFET, respectively.

Figure 3 represents the contour plot of potential with $V_{ds} = 0.0$ V and $V_{gs} = 0.0$ V. The cylindrical dual-metal gate FET has the gate with the lower work function at the drain side, and gate with the higher work function on the other side. The structure parameters of VA-DMCNFET are enlisted in Table 1. The silicon dioxide (SiO_2) is used as the gate dielectric. The doping concentration, N_d, for source(n+)-drain(n+) is $10^{18}/cm^3$. The simulations are run on ATLAS 3D device simulator using the models listed as follows.

MUN and MUP parameters in ATLAS 3D device simulator have been used to set constant values for electron and hole mobilities. The Shockley–Read–Hall (SRH) model is used to simulate the leakage currents. The Auger recombination model is also used to consider minority carrier recombination. Boltzmann model is used for considering the carrier statistics. Lastly, Gummel and Newton techniques have been used for the numerical solutions [22].

Figure 4 shows that the results of the simulated device showcase trends similar to the published research work of Yang et al. [21]. The absolute values, however, cannot be compared owing to minor differences in configuration of the two devices.

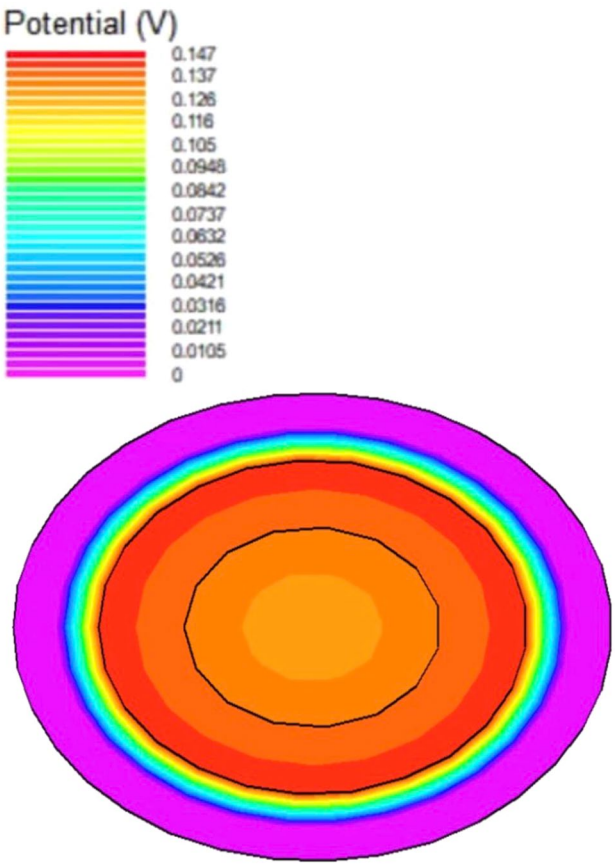


Fig. 3 Potential contour profile of VA-DMCNFET with $V_{ds}=0.0$ V and $V_{gs}=0.0$ V

Table 1 Device parameters

Parameter	Symbol	Value(s)
Channel length (μm)	L	1,2,3 (at fixed channel radius)
Channel radius (nm)	R	2,3,4,5 (at fixed channel length)
Silicon dioxide thickness (nm)	t_{ox}	2
Source/drain length (nm)	L_s/L_d	14
Gate metal work function (eV)	ϕ_1, ϕ_2	4.4, 4.7
Oxide permittivity	ε_{ox}	3.9

3 Results and discussion

The structure and characteristics of the device are developed and extracted, respectively, via SILVACO 3D ATLAS TCAD simulator.

In Fig. 5, the performance of VA-DMCNFET has been compared against NWFET, with both devices having same geometry and constant parameters (L = 100 nm).

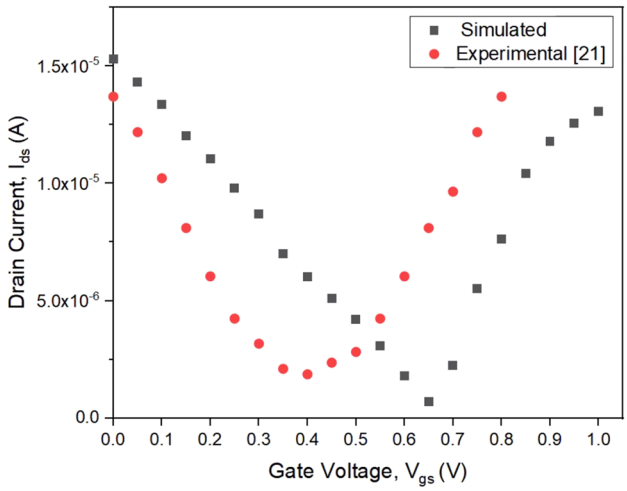


Fig. 4 Calibration of simulation results with experimental data at $V_{ds}=0.8$ V

The results improve drastically with the inclusion of CNT as channel when compared to the conventional silicon channel owing to excellent electrical properties of CNT. Figure 5a shows transfer characteristics of the VA-DMCNFET and NWFET, and it can be seen that VA-DMCNFET exhibits higher ON state current (I_{ds} at $V_{gs}=1.0$ V for constant $V_{ds}=0.5$ V) as compared to NWFET. Furthermore, the VA-DMCNFET exhibits lower channel resistance than NWFET, ensuring high current drive as can be seen from Fig. 5b. It can also be observed that there is achievement of lower threshold voltage in VA-DMCNFET than NWFET. Due to lower threshold voltage, the proposed device gets turned on at a lower gate voltage as compared to NWFET. The extremely high carrier mobility in the CNT in VA-DMCNFET provides high transconductance and operating speed for CNT-based FETs as can be seen from Fig. 5c [23].

The next part of the research intends to draw the focus on the role of plasma parameters on the performance of VA-DMCNFET. Through the means of physical model as shown by Sharma et al. [9], the device geometry and electrical properties of VA-DMCNFET are related to the PECVD parameters. Precise control of the channel radius is a critical predicament for the success of any CNT-based device technology. Channel radius of the CNTFET has a strong impact on the band gap of the semiconducting CNT channel, which further influences various properties of the device [24, 25]. The value of CNT radius decreases with increase in values of plasma parameters: electron density n_{eo} and electron temperature T_{eo}, ion density n_{io} and ion temperature T_{io}. For a given set of plasma parameters, CNT radius increases initially with time, till a saturation value is achieved. As n_{eo} and T_{eo} increase, positively charged ions and electrons are created at the expense of

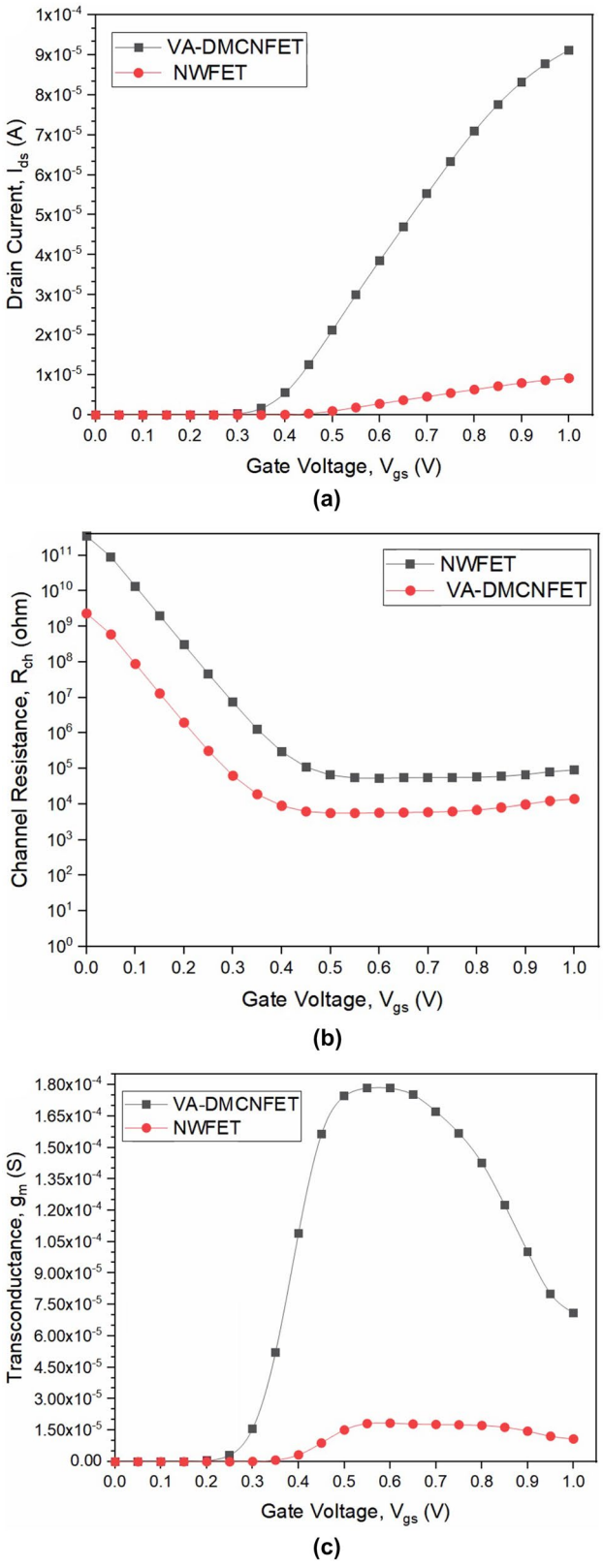


Fig. 5 Comparison between proposed device VA-DMCNFET and NWFET having same parameters: Variation in **a** drain current I_{ds} with gate voltage V_{gs} **b** channel resistance R_{ch} with V_{gs} (in log scale) **c** transconductance g_m with V_{gs} at $V_{ds}=0.5$ V

neutral atoms due to ionization of neutral atoms. Because the accumulation of neutral atoms leads to growth of CNT, the gradual reduction in the density of neutral atoms with increasing values of plasma parameters leads to a decrease in the radius of the CNT [9, 28]. It becomes quintessential to optimize various plasma parameters for optimum device performance, considering they can directly affect the geometry of the synthesized CNT device. Plasma parameters control the CNT channel radius and CNT channel length which further gives them control over the conductivity of the channel, gate controllability, contact resistance and other electrostatics. In this part, we intend to understand how these plasma parameters influence various properties of the device. This research breaks down the variation in plasma parameters into two sets:

Set 1: Variation in plasma parameters corresponding to different values of channel radius for a fixed value of channel length. The channel radius which correlates with the chosen value of plasma parameters varies from 2 to 5 nm (cf. Table 2).

Set 2: Variation in plasma parameters corresponding to different values of channel length for a fixed value of channel radius. The channel length which correlates with the chosen value of plasma parameters varies from 1 to 3 μm (cf. Table 3).

The value of channel radius and channel length pertaining to the chosen plasma parameters is mentioned as follows.

Based on the existing knowledge of Plasma Physics, surface potential is given by

$$V_s = -\frac{2e}{L}\log\left(\frac{\lambda_D}{R}\right) \tag{1}$$

where e = electronic charge, λ_D = Plasma Debye length.

Hence, value of channel radius R can be written as

$$R = \exp\left(\frac{V_s L}{2e}\right)\lambda D,$$

$$Also,\ \lambda_D^{-2} = \lambda_e^{-2} + \lambda_i^{-2} \tag{2}$$

where λ_e = electron plasma Debye length,

λ_i = ion plasma Debye length and can be written as

$$\lambda e = \sqrt{\frac{T_{eo}}{4\pi n_{eo} e^2}}\quad \lambda i = \sqrt{\frac{T_{io}}{4\pi n_{io} e^2}} \tag{3}$$

$$R^2 = \exp\left(\frac{VsL}{e}\right)\left[\frac{\lambda e^2 \lambda i^2}{\lambda e^2 + \lambda i^2}\right] \tag{4}$$

Putting the values of λ_e and λ_i, Eq. (4) can be rewritten as

Table 2 Variation in plasma parameters: n_{eo} and T_{eo}, n_{io} and T_{io} with channel radius (For a fixed value of channel length = 100 nm)

Case	Channel radius (nm)	Plasma parameter 1 (electron density n_{eo} in cm^{-3})	Plasma parameter 2 (electron temperature T_{eo} in eV)	Plasma parameter 3 (ion density n_{io} in cm^{-3})	Plasma parameter 4 (ion temperature T_{io} in Kelvin)
1	2	10^7	0.5	5×10^6	2400
2	3	5×10^6	0.45	4.2×10^6	2300
3	4	10^6	0.4	2.5×10^6	2200
4	5	10^5	0.3	10^6	2100

Table 3 Variation in plasma parameters: n_{eo} and T_{eo}, n_{io} and T_{io}, with channel length (For a fixed value of channel radius = 5 nm)

Case	Channel length (μm)	Plasma parameter 1 (electron density n_{eo} in cm^{-3})	Plasma parameter 2 (electron temperature T_{eo} in eV)	Plasma parameter 3 (ion density n_{io} in cm^{-3})	Plasma parameter 4 (ion temperature T_{io} in Kelvin)
1*	1	5×10^8	1.1	5×10^7	2180
2*	2	10^9	1.2	5×10^8	2210
3*	3	5×10^9	1.3	10^9	2250

$$R^2 = \exp\left(\frac{V_s L}{e}\right)\left[\frac{1}{\frac{4\pi n_{eo}e^2}{T_{eo}} + \frac{4\pi n_{io}e^2}{T_{io}}}\right]$$

or

$$R^2 = \exp\left(\frac{V_s L}{e}\right)\left[\frac{1}{4\pi e^2\left(\frac{n_{eo}}{T_{eo}} + \frac{n_{io}}{T_{io}}\right)}\right] \quad (5)$$

The band gap of CNT channel can be defined as [26]:

$$Eg = \frac{2aoE\pi}{D_{CNT}} \quad (6)$$

where E_π = tight-binding parameter,

a_o = carbon–carbon distance in CNT structure, D_{CNT} = diameter of CNT.

By inserting the value of radius of CNT from Eq. (5), we obtain

$$Eg = aoE\pi \exp\left(-\frac{VsL}{2e}\right)(4\pi e^2)^{1/2}\left(\frac{n_{eo}}{Teo} + \frac{n_{io}}{Tio}\right)^{1/2}$$

$$Eg = k\exp\left(-\frac{VsL}{2e}\right)\left(\frac{neo}{Teo} + \frac{nio}{Tio}\right)^{1/2}, \quad (7)$$

where $k = aoE\pi(4\pi e^2)^{1/2}$

Threshold voltage can be defined as

$$Vth \approx \frac{Eg}{2e} = \frac{aoE\pi}{eD_{CNT}} = \frac{aoE\pi}{2eR} \quad (8)$$

(without considering the DIBL effect)

The drain current in the linear regime is

$$I_{ds} = \frac{\mu C}{L^2} V_{ds}(V_{gs} - V_{th}), \quad (9)$$

C = Capacitance of the CNT and μ = apparent mobility in linear regime.

By substituting the value of R from Eq. (5) in Eqs. (7–9), for the first time, the relation between the above-mentioned properties, e.g., band gap, threshold voltage can be obtained in terms of plasma parameters.

The transfer characteristics for different values of plasma parameters that correspond to different values of channel radius (for a fixed value of channel length) at V_{ds} = 0.5 V are shown in Fig. 6a and for different values of channel length (for a fixed value of channel radius) are shown in Fig. 6b. The output characteristics for different values of plasma parameters that correspond to different values of channel radius (for a fixed value of channel length) are shown in Fig. 7a and for different values of channel length (for a fixed value of channel radius) are shown in Fig. 7b.

On reducing the values of plasma parameters, the value of saturated drain current increases. The semiconducting CNT band gap is inversely proportional to the channel radius as can be seen from Eq. (6) mentioned above. Since I_{ds} of a FET depends on the total charge that fills up the first sub-band, I_{ds} becomes dependent on channel radius and hence on the plasma parameters. Scaling down plasma parameters leads to scaling down of channel length, which further leads

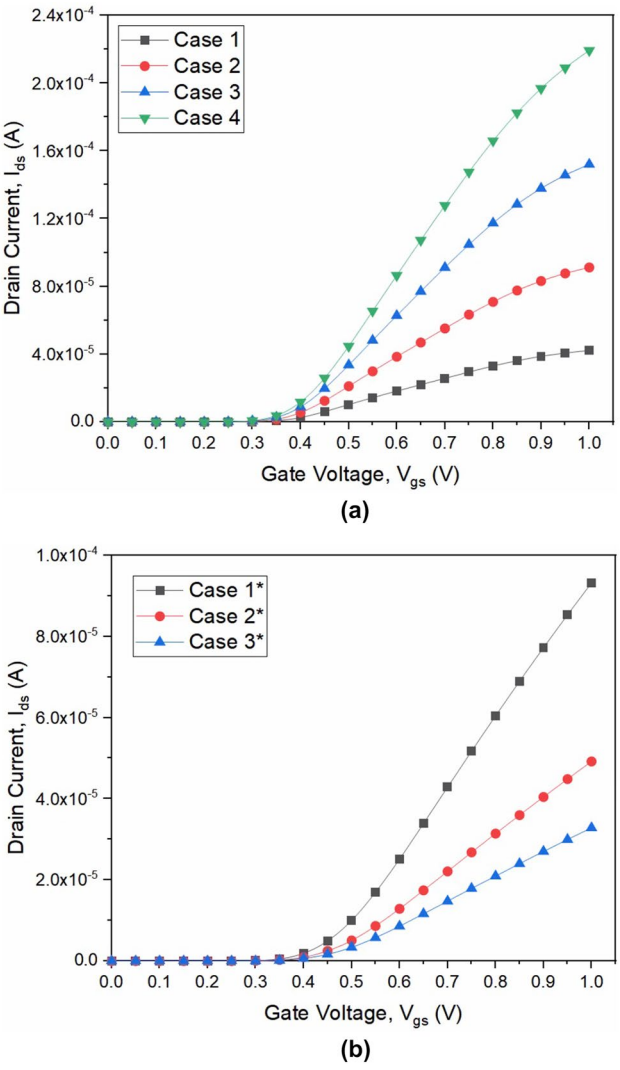


Fig. 6 Transfer characteristics for VA-DMCNFET for different values of plasma parameters corresponding to different values of **a** channel radius (in nm) (Table 2), **b** channel length (in μm) (Table 3) at $V_{ds} = 0.5$ V,

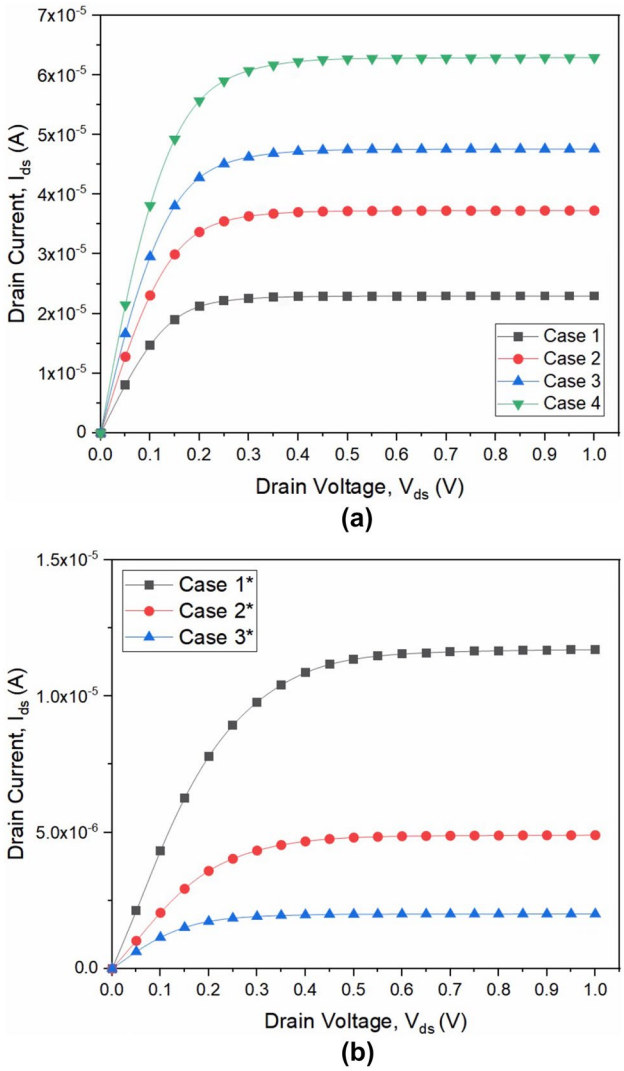


Fig. 7 Output characteristics of VA-DMCNFET for different values of plasma parameters corresponding to different values of (a) channel radius (in nm) (Table 2), **b** channel length (in μm) (Table 3) at $V_{gs} = 0.5$ V

to better gate controllability resulting in increasing the value of saturated drain current. This makes the right balance of plasma parameters extremely important for efficient working of a device.

Figure 8a and b shows the variation in channel resistance R_{ch} with V_{gs} (at fixed value of $V_{ds} = 0.5$ V) for varying values of plasma parameters corresponding to different values of channel radius and channel length, respectively. It can be observed that R_{ch} reduces with decreasing values of plasma parameters, thus ensuring high current drive. From Fig. 9, it can be observed that on reducing the values of plasma parameters corresponding to different values of channel radius, threshold voltage decreases. With reducing plasma parameters that directly affect the channel radius, gate electrode's control on the mid of the channel decreases, leading to reduced electric field lines originating from the gate electrode. This reduces the threshold voltage.

I_{ON}/I_{OFF} current ratio is an essential parameter to measure the digital performance of the device. As can be seen from Fig. 10, with increasing values of plasma parameters, the I_{ON}/I_{OFF} ratio increases. Figure 11 shows the variation in threshold voltage and I_{ON}/I_{OFF} current ratio for different values of plasma parameters corresponding to different values of channel length. The level of ON current increases with reducing plasma parameters. For Fig. 10, with reducing plasma parameters, channel radius increases, which reduces the bandgap and hence increasing ON current [32]. For Fig. 11, with increasing values of plasma parameters, channel length increases which reduces charge carrier density. However, with increasing ON current, OFF state current

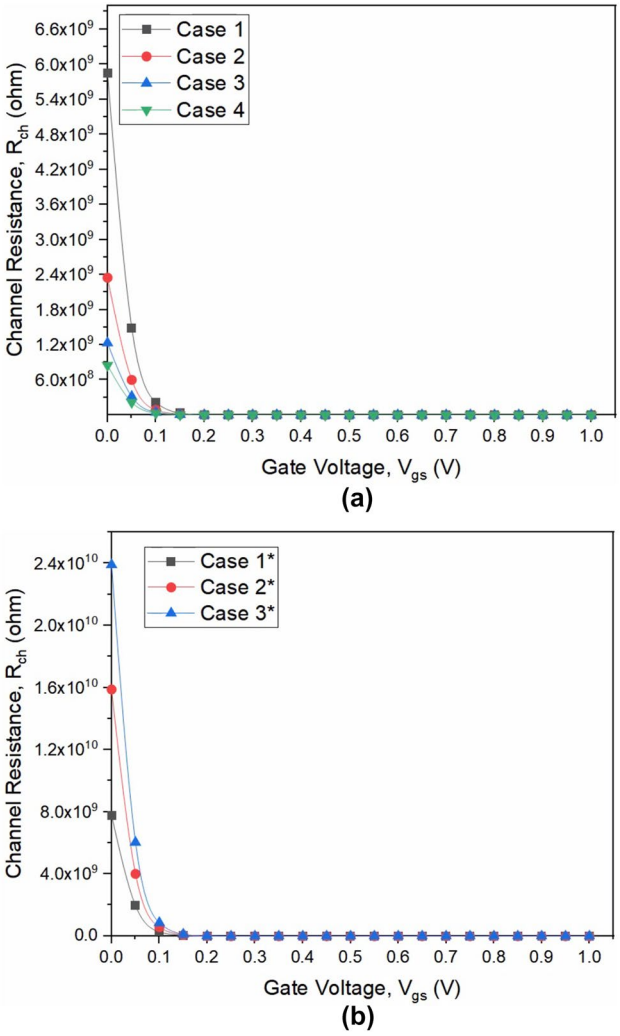


Fig. 8 R_{ch} vs V_{gs} for different values of plasma parameters corresponding to different values of **a** channel radius (in nm) (Table 2), **b** channel length (in μm) (Table 3) at $V_{ds} = 0.5$ V

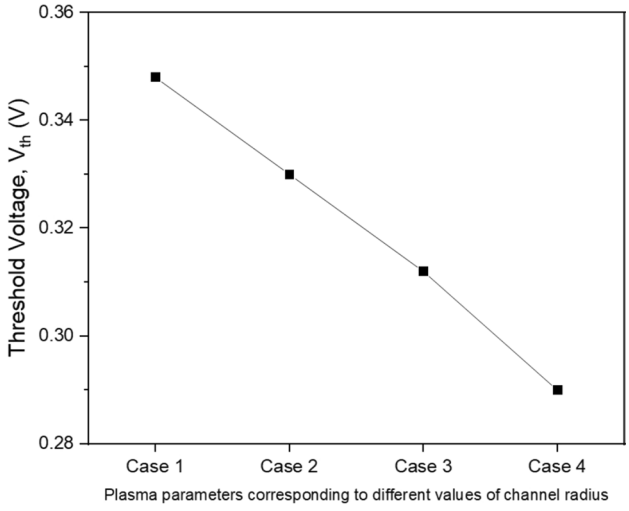


Fig. 9 Variation in V_{th} for different values of plasma parameters corresponding to different values of channel radius (in nm) (Table 2) at $V_{ds} = 0.1$ V

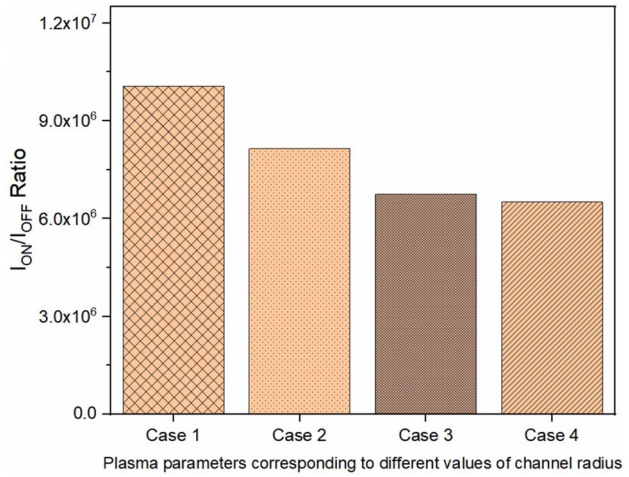


Fig. 10 Variation in I_{ON}/I_{OFF} ratio for different values of plasma parameters corresponding to different values of channel radius (in nm) (Table 2) at $V_{ds} = 0.5$ V

also increases due to leakages, which should be factored in accordingly.

Transconductance (g_m) is a critical property that governs the performance of a FET. It is known that g_m of a device increases with the decrease in band gap owing to a stronger control of gate on the channel. A larger value of g_m is capable of delivering larger gain of device, when all other factors are kept constant. Figure 12a and b shows the variation in g_m with V_{gs} (at fixed value of $V_{ds} = 0.5$ V) for varying values of plasma parameters corresponding to different values of channel radius and channel length, respectively. The value of g_m increases with decreasing values of plasma parameters which can be attributed to higher values of drain current as is observed from Fig. 6a and b. A higher transconductance of the device implies a higher transport efficiency and hence a better applicability for analog devices. Hence, the higher the value of transconductance, the better it is.

Mathematically, output conductance, g_d, is defined as the derivative of I_{ds} to V_{ds} at fixed gate bias V_{gs}. Figure 13a shows g_d vs V_{ds} for different values of plasma parameters corresponding to different channel radius for a fixed channel length at fixed $V_{gs} = 0.5$ V. As observed from Fig. 13a, value of g_d is higher at lower values of plasma parameters corresponding to higher value of channel radius. As the output conductance is a derivative of the drain current, an improvement in the drain current would subsequently result in the increase in output conductance. In Fig. 13b, the value of g_d

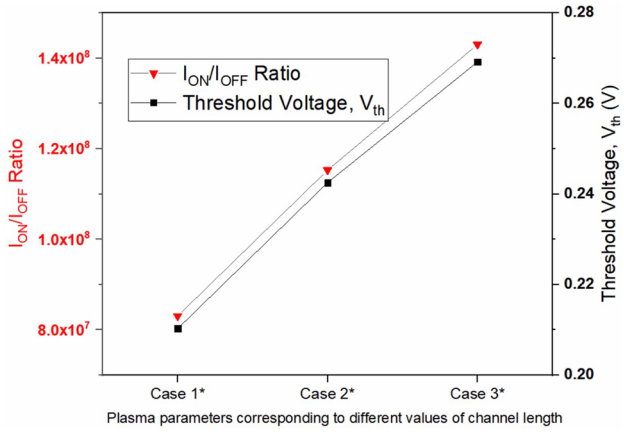


Fig. 11 Variation in threshold voltage, V_{th} and I_{ON}/I_{OFF} ratio for different values of plasma parameters corresponding to different values of channel length (in μm) (Table 3) at $V_{ds} = 0.5$ V

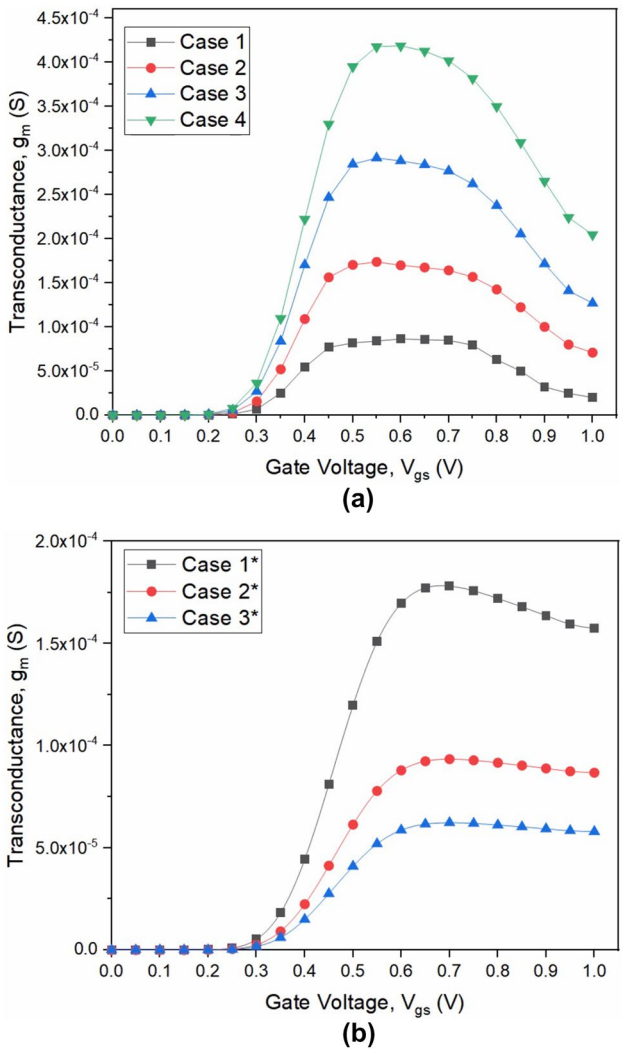


Fig. 12 g_m vs V_{gs} for different values of plasma parameters corresponding to different values of **a** channel radius (in nm) (Table 2) **b** channel length (in μm) (Table 3) at $V_{ds} = 0.5$ V

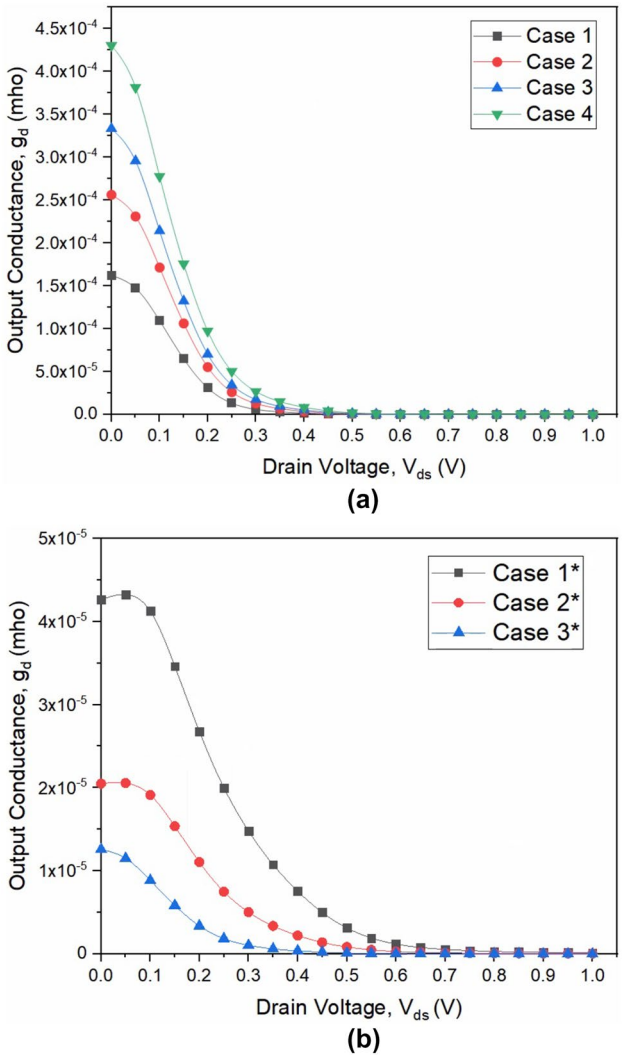


Fig. 13 g_d vs V_{ds} for different values of plasma parameters corresponding to different values of **a** channel radius (in nm) (Table 2) **b** channel length (in μm) (Table 3) at $V_{gs} = 0.5$ V

increases with scaling down of plasma parameters and is attributed to the fact that higher drain currents are observed for lower value of plasma parameters and lower values of channel length due to better gate controllability. Hence, lower plasma parameter conditions are the most appropriate for higher values of g_d.

Early voltage V_{EA} is defined as the ratio of drain current I_{ds} and output conductance g_d, i.e.,

$$V_{EA} = \frac{I_{ds}}{g_d}$$

The gain A_V of the proposed device can be expressed as

$$A_V = \frac{g_m}{I_{ds}} \times V_{EA}$$

Higher values of V_{EA} and A_V are required for better analog performance [33].

Figure 14a and b shows the shift in V_{EA} with V_{ds} (at fixed $V_{gs}=0.5$ V) for different values of plasma parameters corresponding to varying values of channel radius (at fixed channel length) and varying values of channel length (at fixed value of channel radius), respectively. As observed, the value of early voltage increases on increasing the values of the plasma parameters. While both I_{ds} and g_d reduce with increasing plasma parameters, the rate of reduction in g_d is more pronounced in this case, leading to increase in V_{EA}. Higher values of plasma parameters are required for higher V_{EA} and hence higher gain A_v which finds extensive applications for high-gain amplification purposes.

Cutoff frequency f_c can be defined as

$$f_c = \frac{g_m}{2\pi C_{gg}}$$

It directly depends on transconductance g_m and is inversely related to gate capacitance C_{gg} as can be seen above [27]. It is a critical parameter for comparing transistors for RF applications. A large value of f_c is preferred since the high frequency operation of a FET is directed by the same. Figure 15a shows that the value of f_c increases with reduction in plasma parameters (on increasing channel radius) as higher value of transconductance will direct higher value of cutoff frequency.

From Fig. 15b, on reducing the values of plasma parameters (for reducing channel length), the value of f_c increases due to increase in transconductance and reduction in parasitic gate capacitance. Hence, for high frequency

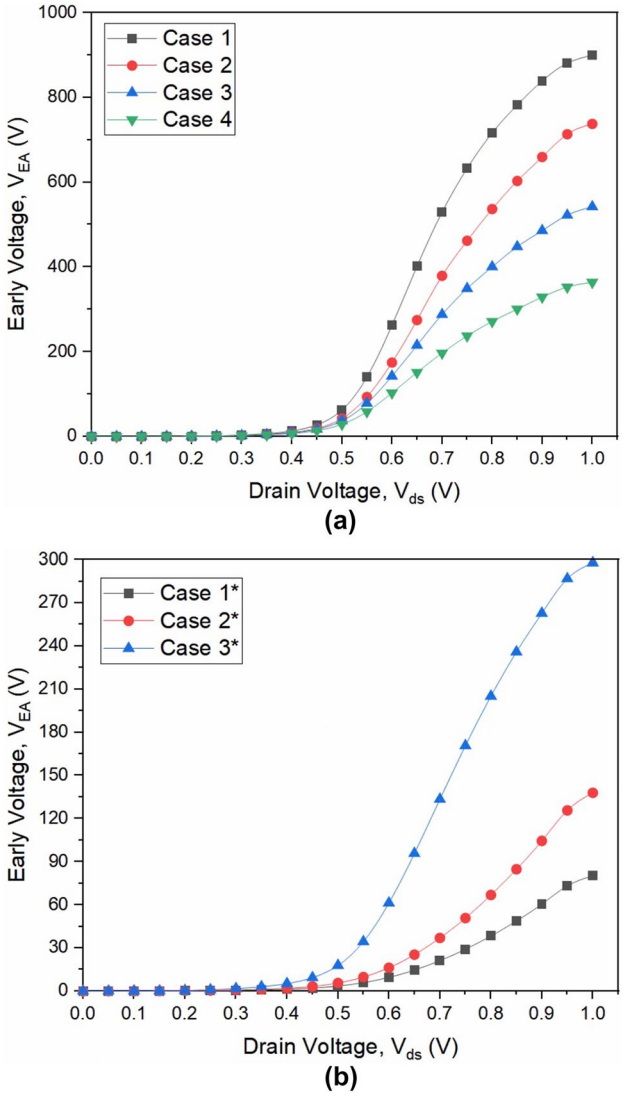


Fig. 14 V_{EA} vs V_{ds} for different values of plasma parameters corresponding to different values of **a** channel radius (in nm) (Table 2) **b** channel length (in μm) (Table 3) at $V_{gs}=0.5$ V

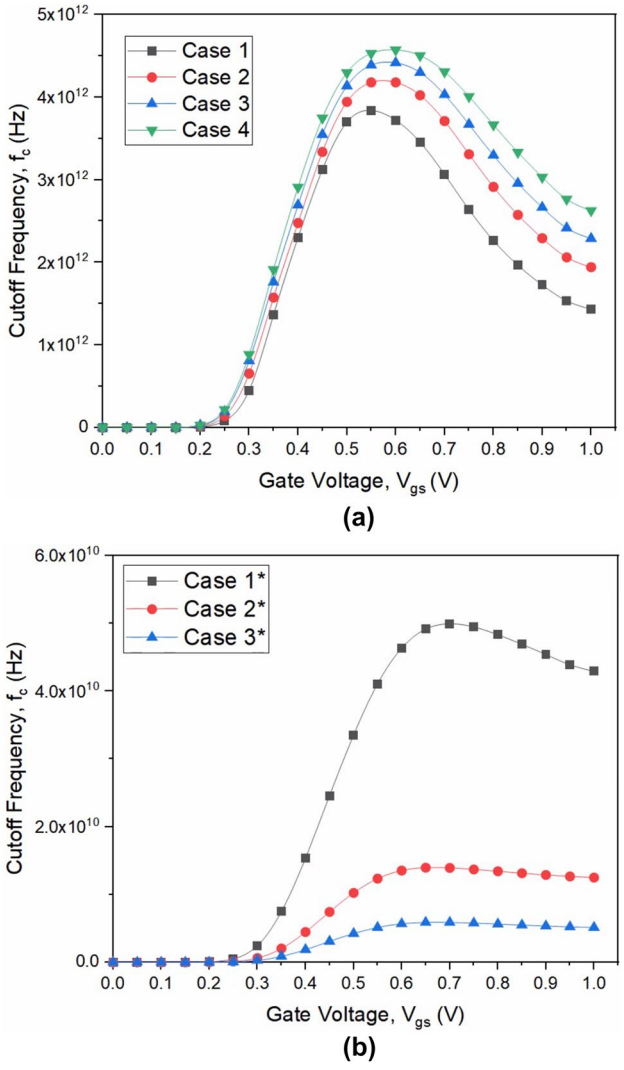


Fig. 15 f_c vs V_{gs} for different values of plasma parameters corresponding to different values of **a** channel radius (in nm) (Table 2) **b** channel length (in μm) (Table 3) at $V_{ds}=0.5$ V

Table 4 Summary table

S. No	Transistor property	Relation with plasma parameters (n_{eo}, T_{eo} & n_{io}, T_{io})
1	Drain current	Increases with decreasing values of plasma parameters
2	Channel resistance	Decreases with decreasing values of plasma parameters
3	Threshold voltage	Decreases with decreasing values of plasma parameters
4	I_{ON}/I_{OFF} ratio	Decreases with decreasing values of plasma parameters
5	Transconductance	Increases with decreasing values of plasma parameters
6	Output conductance	Increases with decreasing values of plasma parameters
7	Early voltage	Decreases with decreasing values of plasma parameters
8	Cutoff frequency	Increases with decreasing values of plasma parameters

operation of the device, lower values of plasma parameters are required.

The summary of relation between plasma parameters and transistor performance is given in Table 4.

The simulated device VA-DMCNFET has a wide range of applications. Plasma-assisted devices find their usage in biosensors, logic gates, etc. Due to high values of drive current and transconductance, the proposed device can be used to maximize speed for digital logic applications [29]. CNTs are ideal for on-chip interconnect applications in microelectronics since they can carry high current densities with fixed resistance over several micrometers. [30–32]. The semiconducting CNTs are responsive to their environment and hence used in bio- and electro-chemical sensors for detecting biological species such as proteins, DNA, and antibodies [34].

4 Conclusion

The proposed device VA-DMCNFET with vertically aligned semiconducting CNT grown in presence of plasma sheath through the process of PECVD technique as the channel was successfully implemented. The obtained results were compared with the existing experimental observations and are found in good agreement. The proposed device shows far improved metrics as compared to NWFET. Further, using this research, we can conclude that with reducing plasma parameters, i.e., increasing values of CNT channel radius for fixed CNT channel length or decreasing values of channel length for fixed channel radius, the absolute value of saturated drain current, transconductance, output conductance increase whereas channel resistance along with threshold voltage can be seen reducing. Along the same lines, I_{ON}/I_{OFF} current ratio, early voltage and gain of the device increase on increasing the values of plasma parameters. In both the sets, reducing plasma parameters helps in increasing the value of cutoff frequency, thus leading to better RF performance. The research that we have conducted can be used as a base in future by several experimentalists for their respective study. Moreover, our simulation results can be useful for explaining many more experimental findings such as their application in digital devices and biosensors.

Acknowledgements Author Mansha Kansal is thankful to Dr. Sonam Rewari for valuable discussions.

Author contributions Not Applicable.

Funding No funds, grants, or other support was received.

Availability of data and material The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. A.K. Gupta, A. Raman, Performance analysis of electrostatic plasma-based dopingless nanotube TFET. Appl. Phys. A **126**, 573 (2020). https://doi.org/10.1007/s00339-020-03736-7
2. H.M. Fahad, M.M. Hussain, High-performance silicon nanotube tunneling fet for ultralow-power logic applications. IEEE Trans. Electron Devices **60**(3), 1034 (2013). https://doi.org/10.1109/TED.2013.2243151
3. C. Qiu, Z. Zhang, M. Xiao, Y. Yang, D. Zhong, L.-M. Peng, Scaling carbon nanotube complementary transistors to 5-nm gate lengths. Science **355**(6322), 271 (2017)
4. I. Santolia, A. Tewari, S.C. Sharma, R. Sharma, Effect of doping on growth and field emission properties of spherical carbon nanotube tip placed over cylindrical surface. Phys. Plasmas **21**(6), 063508 (2014)
5. A. Rochefort, D.R. Salahub, P. Avouris, Effects of finite length on the electronic structure of carbon nanotubes. J. Phys. Chem. B **103**(4), 641–646 (1999)
6. T. Mizutani, H. Ohnaka, Y. Okigawa, S. Kishimoto, Y. Ohno, A study of preferential growth of carbon nanotubes with semiconducting behavior grown by plasma-enhanced chemical vapor deposition. J. Appl. Phys. **106**(7), 073705 (2009). https://doi.org/10.1063/1.3234389
7. D. Phokharatkul, Y. Ohno, H. Nakano, S. Kishimoto, T. Mizutani, High-density horizontally aligned growth of carbon nanotubes

with Co nanoparticles deposited by arc-discharge plasma method. Appl. Phys. Lett. **93**(5), 053112 (2008)
8. Y. Ono, S. Kishimoto, Y. Ohno, T. Mizutani, Thin film transistors using PECVD-grown carbon nanotubes. Nanotechnology **21**(20), 205202 (2010)
9. S.C. Sharma, A. Tewari, Effect of plasma parameters on growth and field emission of electrons from cylindrical metallic carbon nanotube surfaces. Phys. Plasmas **18**(8), 083503 (2011)
10. A. Tewari, S.C. Sharma, Theoretical modeling of temperature dependent catalyst-assisted growth of conical carbon nanotube tip by plasma enhanced chemical vapor deposition process. Phys. Plasmas **22**(2), 023505 (2015)
11. T. Kato, R. Hatakeyama (2010), Growth of single-walled carbon nanotubes by plasma CVD. J. Nanotechnol.
12. C. Suresh, N. Sharma, Phys. Plasmas **22**(12), 123517 (2015)
13. R. Gupta, S.C. Sharma, R. Sharma, Mechanisms of plasma-assisted catalyzed growth of carbon nanofibres: a theoretical modelling. Plasma Sources Sci Technol **26**(2), 024006 (2017)
14. Y.C. Choi, D.J. Bae, Y.H. Lee, B.S. Lee, G.-S. Park, W.B. Choi, J.M. Kim, Growth of carbon nanotubes by microwave plasma-enhanced chemical vapor deposition at low temperature. J. Vac. Sci. Technol. A: Vac., Surf., Films **18**(4), 1864–1868 (2000)
15. L. Nougaret, H. Happy, G. Dambrine, V. Derycke, J.-P. Bourgoin et al., 80 GHz field-effect transistors produced using high purity semiconducting single-walled carbon nanotubes. Appl. Phys. Lett. **94**, 243505 (2009)
16. S. Jing Guo, A. Hasan, G.B. Javey, M. Lundstrom, Assessment of high-frequency performance potential of carbon nanotube transistors. IEEE Trans. Nanotechnol. **4**(6), 715–721 (2005). https://doi.org/10.1109/TNANO.2005.858601
17. Y. Li, D. Mann, M. Rolandi, W. Kim, A. Ural, S. Hung, A. Javey, J. Cao, D. Wang, E. Yenilmez, Q. Wang, J.F. Gibbons, Y. Nishi, H. Dai, Preferential growth of semiconducting single-walled carbon nanotubes by a plasma enhanced CVD method. Nano Lett. **4**(2), 317–321 (2004)
18. T. Mizutani, Y. Ohno, S. Kishimoto, Electrical properties of carbon nanotube FETs. 2008 International Conference on Advanced Semiconductor Devices and Microsystems, pp. 1–8,(2008).doi: https://doi.org/10.1109/ASDAM.2008.4743290
19. F. Bashir, S.A. Loan, M. Rafat, A.R.M. Alamoud, S.A. Abbasi, A high-performance source engineered charge plasma-based schottky MOSFET on SOI. IEEE Trans. on Electron Devices **62**(10), 3357 (2015)
20. S.K. Srivastava, A.K. Shukla, V.D. Vankar, V. Kumar, Growth, structure and field emission characteristics of petal like carbon nano-structured thin films. Thin Solid Films **492**(1–2), 124–130 (2005). https://doi.org/10.1016/j.tsf.2005.07.283
21. X. Yang, K. Mohanram, Modeling and performance investigation of the double-gate carbon nanotube transistor. IEEE Electron Device Lett. **32**(3), 231–233 (2011). https://doi.org/10.1109/LED.2010.2095826
22. ATLAS User's Guide, SILVACO Int., Santa Clara, CA, USA, Version 5.26.1.R; (2018)
23. L. Liu, C. Qiu, D. Zhong, J. Si, Z. Zhang, L.-M. Peng, Scaling down contact length in complementary carbon nanotube field-effect transistors. Nanoscale **9**(27), 9615–9621 (2017)
24. G. Jo, J. Maeng, T.-W. Kim, W.-K. Hong, B.-S. Choi, T. Lee, Channel-length and gate-bias dependence of contact resistance and mobility for In2O3 nanowire field effect transistors. J. Appl. Phys. **102**(8), 084508 (2007)
25. R. Martel, T. Schmidt, H.R. Shea, T. Hertel, Ph. Avouris, Single- and multi-wall carbon nanotube field-effect transistors. Appl. Phys. Lett. **73**, 2447 (1998)
26. M.H. Moaiyeri, F. Razi, Performance analysis and enhancement of 10-nm GAA CNTFET-based circuits in the presence of CNT-metal contact resistance. J Comput. Electron. **16**, 240–252 (2017)
27. S.M. Sze, K.K. Ng, *Physics of Semiconductor Devices* (Wiley, New York, 2006)
28. K Ostrikov Xu Shuyan P-A Nanofabrication (2007). From Plasma Sources to Nanoassembly, Wiley-VCH Germany https://doi.org/10.1002/9783527611553
29. G.J. Brady, A.J. Way, N.S. Safron, H.T. Evensen, P. Gopalan, M.S. Arnold, Quasi-ballistic carbon nanotube array transistors with current density exceeding Si and GaAs. Sci. Adv. **2**(9), e1601240–e1601240 (2016). https://doi.org/10.1126/sciadv.1601240
30. A. Graham, G. Duesberg, W. Hoenlein et al., How do carbon nanotubes fit into the semiconductor roadmap? Appl. Phys. A **80**, 1141–1151 (2005)
31. A.D. Franklin, S.O. Koswatta, D.B. Farmer, J.T. Smith, L. Gignac, C.M. Breslin, J. Tersoff et al., Carbon nanotube complementary wrap-gate transistors. Nano Lett. **13**(6), 2490–2495 (2013). https://doi.org/10.1021/nl400544q
32. T. Dang, L. Anghel, R. Leveugle, CNTFET basics and simulation. Int. Conf. Des. Test Integr. Syst. Nanoscale Technol., DTIS **2006**, 28–33 (2006). https://doi.org/10.1109/DTIS.2006.1708731
33. S. Rewari, V. Nath, S. Haldar, S.S. Deswal, R.S. Gupta, Improved analog and AC performance with increased noise immunity using nanotube junctionless field effect transistor (NJLFET). Appl. Phys. A (2016). https://doi.org/10.1007/s00339-016-0583-9
34. S.S. Alabsi, A.Y. Ahmed, J.O. Dennis, M.H.M. Khir, A.S. Algamili, A review of carbon nanotubes field effect-based biosensors. IEEE Access **8**, 69509–69521 (2020)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)
Pune, India. Oct 29-30, 2021

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON) | 978-1-6654-2503-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/SMARTGENCON51891.2021.9645749

Retinal Vessel Detection Using Residual Y-net

Anindita Roy
Department of Information Technology
Delhi Technological University
Delhi, India
auni.roy97@gmail.com

Kapil Sharma
Department of Information Technology
Delhi Technological University
Delhi, India
kapil@ieee.org

***Abstract*— In this Corona Pandemic, Diabetic patients are affected a lot. Unfortunately, due to the consumption of steroids, people are affected by mucormycosis, which is a kind of fungal infection, making this situation worse. Patients get swollen red eyes, and diabetic patients are more vulnerable to it, as they suffer from Diabetic Retinopathy. It has become essential to determine the damage caused to the eye to save patients from vision loss. Only doctors can identify how the condition of the eye by physical examination. But this is a tricky and time-consuming job. With the help of fundus photography and deep learning algorithms, the detection and classification process will speed up. There are many existing image detection algorithms, but they do not have efficient feature retention and lightweight architecture model. This paper proposes Residual Y-net architecture that works excellently on a balanced medium-size. With the help of segmented features, it acquires reliable features which help in classification. It is a very lightweight architecture inspired by U-net, Deep Residual U-net, and Y-net. The addition of residual units in the network has significantly improved the accuracy rate. It is observed that a balanced dataset gives a much accurate performance than an unbalanced dataset. The proposed model's test accuracies on medium-size unbalanced and balanced datasets are 90.39% and 93.60%, respectively.**

Keywords*— DR – Diabetic Retinopathy, CNN – Convolutional Neural Network, Res U-net Residual U-net, Res Y-net – Residual Y-net

I. Introduction

Diabetes has become an underlying condition in patients who are affected by the Coronavirus pandemic. Studies show that critical forms of COVID are seen in uncontrolled diabetic patients [1]. The process of healing in their body is slower than other patients that put them at higher risk. [2]Diabetic patients are mostly admitted to ICU (Intensive Care Unit) since they are most likely to have other health critical health conditions like hypertension, cardiovascular disease, etc. making them more vulnerable in this pandemic [3] [4] [5]. The most number of deaths are seen in the old age group with diabetes. However, various medications especially steroids made the situation worse and now people are struck with a fungal infection called, Mucormycosis [6]. Generally, patients are facing swollen red eyes. It is a dangerous infection taking many lives. It is again dangerous for diabetic patients as they suffer from Diabetic Retinopathy. A patient suffered from diabetic ketoacidosis [7]with right eye ptosis. Histopathology results disclosed the patient had mucormycosis. Even while treating for mucormycosis, the patient already [3]had diabetes. Fundus photography showed the presence of non-proliferative diabetic retinopathy. The medications used for Covid treatment are ocular toxic [8]. Few of them are responsible for inducing retinopathy in the patients. Therefore, the detection of diabetic retinopathy is still a critical situation for the medical industry.

Retinopathy results in swelling of blood vessels in the retina of the eyes. It gives mild issues in the vision. But, if not cured early, the patient may lose his or her eyesight [9]. This situation of the eyes of the patient is coined as Diabetic Retinopathy. Harm to veins in the retina happens when the veins get hindered because of high sugar levels in the blood. New blood vessels attempt to develop but are not able to develop appropriately. These create small lumps in the blood vessels, which may now and then hole blood into the retina [10]. The width of the veins changes and the state of the retina gets highly severe. There are various classes in DR as shown in Fig. 1. Ophthalmologists have specified different forms of damages caused by the retina of the eyes.

1) Microaneurysms [11]*:* These are the red spots that happen in the eye chunk of diabetic patients. Even though they don't influence eye vision, they help in distinguishing the beginning period of diabetes.

2) Cotton-fleece spots [11]*:* They are white and feathery fixes in the eye. These appear when the eye nerve strands damage. There are associated impacts, too like, haemorrhages and microvascular infarcts.

3) Hemorrhages [11]*:* This shows up as splendid red or dim red fixed in the eye. They can seep inside in the eye. The dividers of the veins may break in the retina.

4) Exudates [11]*:* These happen on the outer side of the retina. They show up as sparkly yellowish in shading. They have sharp edges. Indeed, even microaneurysms, which have no bloodline in them, are additionally called hard exudates.

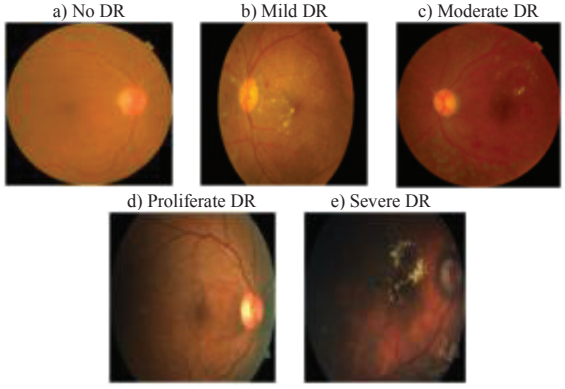


Fig. 1 Sample images of DR classes

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:49:20 UTC from IEEE Xplore. Restrictions apply.

The model might under-fit when a model runs on a small dataset. While in large datasets, the unbalanced class distribution causes the model to overfit. The model will undergo training securely if the dataset is present in similar class distribution, i.e., the number of images in each label is almost the same. U-net [12] is a type of CNN specifically invented for biomedical image classification. It works fine when the dataset is small. The architecture helps to train end to end, and it classifies each pixel in the image into a class label. It is better than patch-wise classification because it is redundant and results in loss of information. In contrast, pixels help in better label classification because it is not redundant.

U-net, Deep Res U-net [13], and Y-net [14] inspired the proposed model. The VGG network serves as the backbone for U-net and Y-net, whereas Res-Net is the backbone of Deep Res U-net. U-net has convolutional blocks that contain a stack of convolution and pooling layers that are responsible for the degradation of feature details. To rectify this problem, Deep Res U-net was proposed which used Residual units [15]. It uplifts the accuracy of neural network. The residual unit is a combination of the input layer, convolutional layers, batch normalization, and activation functions. From each passing layer in a residual unit, the neurons tend to lose a few of the features of the actual layer, so when the original input layer is added with the output layer, then all the lost information along with the processed output can be obtained. To solve the same problem of feature loss, instead of using conventional encoder-decoder architecture, Y-net applies two encoders and one decoder. It adds the output of the convolutional block of the first encoder with the output of the corresponding convolutional block of the second encoder. The proposed model is termed as Deep Res Y-net because it uses residual units [15], and has Y-net architecture. The primary purpose of this method is to contain as much information as possible. It reduces the possibility of degradation of the model, which generally happens when going from higher to lower levels. From each block of the encoders, features are combined and moved to the appropriate stage of the decoder. This help in filling the gaps when any information from the neurons is corrupted or lost at any point.

This paper has five sections. The first section is an introduction to the work, which explains the context of the project, the proposed methodology and dataset information. The second section is a glimpse of work that has been done by various authors in the past years on Diabetic Retinopathy using deep learning architectures, about the dataset they used and results they obtained. The third second goes over U-net, Deep Residual U-net, Y-net, and the proposed Deep Residual Y-net model in great detail. The fourth section explains the dataset, experimental setup to execute the model, and results in graphical forms observed after execution. The final section concludes this paper on its proposed method and future work.

II. Related Works

There has been a considerable amount of work in diabetic retinopathy classification. Many deep learning methods are applied. This paper [16] uses DenseNet over 3050 images obtained from Kaggle. It gives a validation accuracy of 84.10%. It describes how a skewed dataset overfits the model and uses the kappa score to measure the model performance due to the hardware constraints and skewness of the dataset. Integrated Shallow Convolutional Neural Network, on datasets ranging from 200 to 35000 photos in quantity, was used [17]. An integrated system helps to bring out the best part from different neural architectures. This [18] paper uses an ensemble learning approach. It boasts strong base learners to put a competent system forward. It executes 1200 group experiments, in which ensemble learning excels. Influenced by this, many authors integrated various deep learning models to perform image detection. In the following proposed architecture [19], a combination of three convolutional neural architectures (SE- ResNeXt50, EfficientNet-B4, and EfficientNet-B5) is used, then applied to that transfer learning. APTOS2019 Kaggle dataset is used in this paper. It contains 3662 training images in similar class distribution. 0.8265 of accuracy obtained from this method. MobileNet V2 is a deep CNN that has been designed especially for computer vision and smartphone applications. It is a lightweight network. [20] Upon this, fine-tuning is applied, due to which accuracy reaches 91%. The dataset comes from Kaggle, which has 3662 images in total.

U-net is also used for image classifications. For nuclei detection, this paper [21] compared two commonly used CNN methods (CNN1 and CNN2) and Residual inception channel attention(RIC) U-net on Cancer Genomic Atlas dataset. Compared with original U-net and RIC-U-net, they got f1 score of 0.8155 and 0.8278. The In the paper [22], weighted residual U-net has been used in which skip connections are utilized. The method adds the input weights into the network at every level to extract features. The experiment is performed over DRIVE and STARE datasets. The accuracy seized in this process is 0.9655 on the DRIVE dataset, and 0.9693 on the STARE dataset. On the same dataset, [23] the authors in this paper proposed a bridge-style U-net. The approach is to add the nearest feature output from one block into the next block as its input information. [24] [25]The closest features rapidly help in gaining the most relevant information. DRIVE, STARE, and TONGREN datasets were used to train and test the proposed model, where 0.9567, 0.9658, and 0.9652 of accuracy had been recorded over the three datasets, respectively. The spatial attention module is designed for detection and classification, is a part of the convolutional block attention module [26]. It is used [27] to improve the performance of the U-net. This module calculates a spatial attention map using the spatial features. Then, it applies global and average pooling layers over the channel axis to produce a defining quality. It is tested over DRIVE and CHASE DB1 dataset in 100 epochs reaching an accuracy of 0.9618 and 0.9905, respectively. [28] This paper used the SLIC superpixel algorithm, which is a patch generating algorithm, to solve the problem of an imbalanced dataset. In U-net, the authors applied residual connections and inception modules on the IDRiD dataset and obtained 97.95% accuracy. [29]CAR-Unet is introduced in which a modified channel attention network is developed. An updated channel focus network replaces the original blocks. This approach employs layers of max pooling and average pooling. These modifications in U-net helped to get an accuracy of 0.9699, 0.9751, and 0.9743 on DRIVE, STARE, and CHASE datasets, respectively. There are combinations of transfer learning

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:49:20 UTC from IEEE Xplore. Restrictions apply.

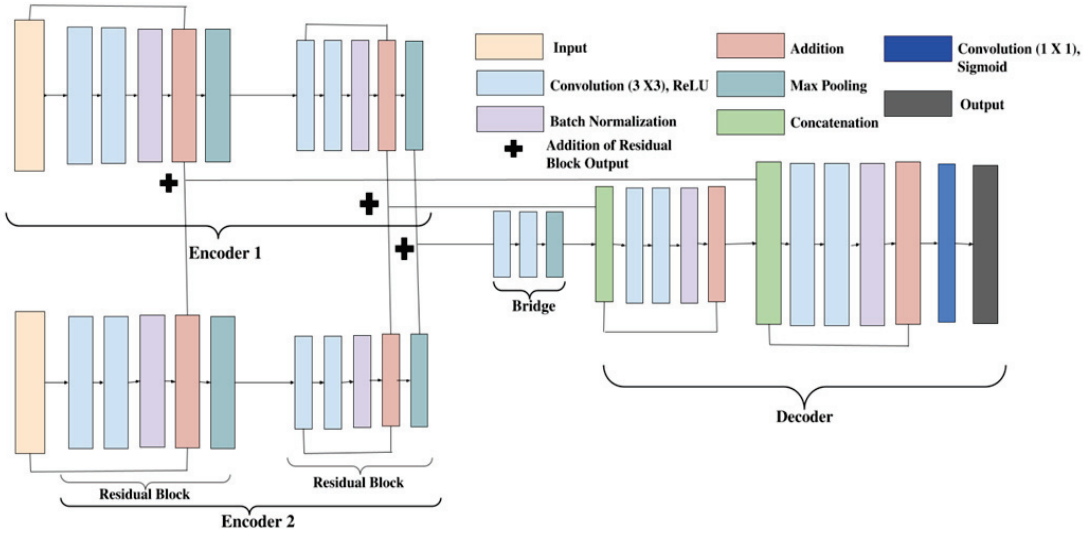


Fig. 2 Residual Y-net Architecture

architectures with U-net. In this paper [30], the authors replaced the usual convolutional blocks of the U-net with the VGG16 network in the encoder part. Instead of using completely connected layers, the Global Max-pooling layer is put after the 5th convolutional layer. Instead of a sigmoid activation function, the last layer in the decoders uses a softmax activation function.

The majority of the research related to the U-net and its modifications are applied on the DRIVE, STARE, and CHASE datasets, which are tiny dataset, as can be seen from the previous found works. The experiment in this paper tries to research upon a much bigger dataset.

Y-net is also recently used in various fields of medical imaging. Y-net [14] has been applied over breast biopsy images for segmentation and classification. It outperforms the traditional encoder-decoder structure by 6%. It based upon the VGG11 encoder. [31]It is added with augmentation to 1000 chest x-ray images. The results were based upon the acceptable percentage that came out to be 95.8%. Then, this architecture has experimented upon polyp detection [32]. It gets compared with U-net and pre-trained U-net. It gives an 84.4% of recall and 85.9% of f1 score, which is way better than the other two models. Although, the precision it obtained is 87.4, which is less than U-net and pre-trained U-net. There is someplace for improvement in the architecture because it performs well on few classification parameters.

III. Methodology

A. Review of U-net

U- Net is introduced in 2015 [12], especially for processing biomedical images. Its basic foundation is upon CNN's conventional architecture. The general convolution neural network is programmed to classify images, where the image is the input and the output is the desired label. U-Net helps to distinguish images pixel by pixel and apply classification over it. It locates the areas of abnormalities. Input and output share the same sizes. U- Net is a kind of architecture that is in U shape. It has two parts that are the left and right sides. The left side, which is called the contracting path, is the usual convolution neural network. The ride side, which is called expansive path, is there so that up-sampling can be done. It is a stack of deconvolution blocks. Since, it extracts semantic feature, therefore, it is able to do better image classification. The issue in U-net is that the input compress, which creates a bottleneck where all features are not passed into the decoder.

B. Review Residual U-Net

Since Residual U-net forwards the input layer information, it is referred to as weighted Res-Unet in many research papers [13]. It is a combination of U-net and residual units.

Stacked residual units are applied successively in residual networks [33]. Every unit is expressed in the following form:

$$k = y(x) + f(x, w) \tag{1}$$

$$xi = A(k) \tag{2}$$

where x is the input, xi is the output of the I-th residual unit, $f(x, y)$ denotes residual function, $A(k)$ denotes the activation function, y(x) gives the identity mapping. Numerous fusion of convolutional layers, ReLu activation function and batch normalization are present in a single residual unit.

It provides two advantages. First, it helps in reducing the computations of training. Second, it propagates information without degrading the model between low and high levels. Both of the advantages are achieved with fewer parameters. Meanwhile, it attains the preferable quality of performance.

C. Review Y-net

Y-net [14] is also a three-part procedure, just like U-net. It has three paths: encoding, bridge, and decoding. It uses two encoders for instance-level segmentation. The input layer for both encoders is the same and the output of two convolutional blocks from the first and second encoders are added and then transferred to the corresponding stage of the decoder block. It helps to keep together features obtained in the condensed form in a much secure way. The decoding goes the same as the U-net decoding block. Convolutional transpose of the input layer

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:49:20 UTC from IEEE Xplore. Restrictions apply.

is done. It is then added to the previously obtained added result of the corresponding outputs from the two encoders' convolutional blocks Different convolutional blocks can be used at any level without affecting the architecture of the network. It proves to be flexible and more combinations of parameters can be applied according to the requirement of the dataset and better performance of the model.

D. Proposed Method: Residual Y-Net

The proposed method is the combination of both Residual Unit and Y-net as shown in Fig. 2. The architecture is in the form of Y-net applying residual unit in the encoding and decoding block. It serves quality performance. At every level, the features are added from every other encoding level, therefore there is less degradation in the model. Just like Y-net, it has two encoders, one bridge and one decoder. Each encoder has two residual units. Every residual unit contains two convolutional layers with 3 X 3 kernel size and ReLU activation function, a batch normalization layer, and a dropout layer. After obtaining Layered Output, a 2 X 2 max-pooling layer with a stride equal to two is applied in the block for down-sampling. This Layered Output is added with its corresponding Layered Output from the second encoder to retain the best segmentation features and label in the best possible way, producing the Block Encoded Output. Just like in Residual U-net, it helps to keep features obtained in the condensed form. The difference is that this Block Encoded Output of Residual Y-net contains double number features. It increases the possibility of better label classification. Thebridge part has two 3 X 3 convolutional layers in which the addition of the output from the two encoders is passed on as input. The decoding part is the same as Y-net. The Decoded Output is generated by concatenating the Convolutional Transpose of the input with Block Encoded Output from corresponding levels. It is followed by two convolutional layers having 3 X 3 kernel size and RelU activation function, a batch normalization layer. The obtained output from the block is again concatenated with Decoded Output of the current block.

The result of the last residual block of the decoder will be a multi-channel feature map and to perform classification over the feature map, a dense layer has been added. The size of the dense layer output is equal to the number of classes in the dataset.

IV. Experiment and Results

A. Dataset

The datasets are taken from APTOS 2019 [34] blindness detection competition which took place on Kaggle. The datasets are classified into five categories, which are No_DR, Mild, Moderate, Severe and Proliferate DR. The total numbers of images are 13000. The competition dataset is unbalanced. From that dataset, Dataset1 which is unbalanced (Fig. 3) and Dataset2 which is a balanced dataset are derived to observe the performance of the proposed algorithm on both datasets and it consists of 3662 and 4918 images respectively. As in the bar graph can be seen (Fig. 4), the number of pictures in each label is quite balanced, i.e. has a similar distribution of data among the classes, so the chances of random division into training and test set will bring better diversity.

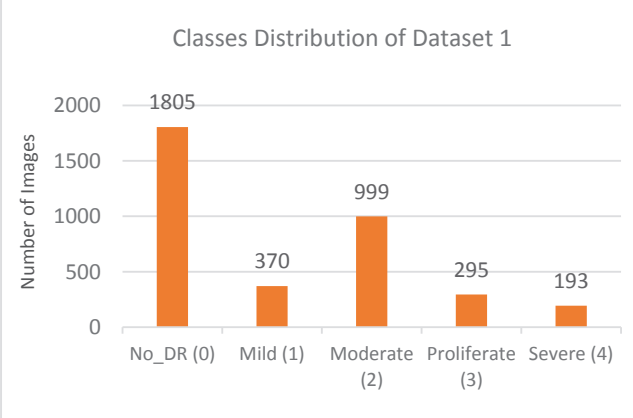


Fig. 3 Classes Distribution of Dataset 1

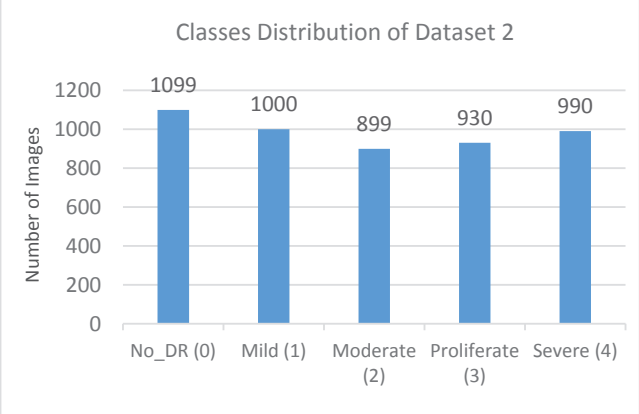


Fig. 4 Classes Distribution of Dataset 2

B. Experimental Setup

The experiment was implemented on a Keras v2.0 framework and trained on NVIDIA Tesla K80 GPU with 12 GB of memory. The images are augmented. Image Augmentation increases the amount and diverseness of data during the scarcity of publicly available data. In image augmentation, the raw input images are rotated, sheared, clipped, zoomed, flipped, and contrast is changed. The APIs have an Image Data Generator, in which various parameters are put. The rotation range, shearing range, and other parameters can be specified and passed to Image Data Generator according to the requirements. Adam optimizer has been used in all the networks with categorical cross-entropy as a loss function. The training and testing data is split into 80-20 ratio randomly.

C. Results

The network's requirements decide the number of epochs. More blocks will result in the saturation on the image condensation since the size of the image is small. So basically, the independent learning of individual neurons decreases. After adding the required number of dropouts, performance gets better. In proposed method which is Residual Y-net, residual units are added, which helped in increasing the high resolution information (segmented results). The parameters increased as compared to Y-net, U-net and Deep Residual U-net. But, there is significant decrease in training epochs. It takes 50 epochs to

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:49:20 UTC from IEEE Xplore. Restrictions apply.

train on Dataset 1(Fig. 5). The Training accuracy is 98.92%, and test accuracy is 90.39%. On Dataset 2, within nine epochs, the test accuracy reaches 93.60%, though the training accuracy is 98.65% (Fig. 6).

In U-net, the training time or the number of epochs also decreases. In 70 epochs, Dataset 1 records train and test accuracies of 96.07% and 86.78%, respectively. While on Dataset 2, in 25 epochs, the training accuracy reaches 99.59%. The Test accuracy comes out to be 89.69%. The setup is kept the same for Y-net. Only the number of encoders is two. At the end of the residual block, the parallel residual blocks from other encoder are added for feature map collection. The number of epochs was high to watch the convergence in training accuracy. Dataset 1 receives 99.82% of train accuracy and 88.75% of test accuracy. Similarly in Y-net, within 25 epochs the training accuracy reaches up to 99.06%, but the test accuracy goes up to 90.57%. It gives slightly better performance in the neural network due to feature retention. Residual U-net is implemented by replacing the convolutional blocks with residual blocks in U-net. A subsequent increase in accuracy is adding residual units. Over Dataset 1, training and test accuracy reach 98.67% and 89.10%, respectively. On Dataset 2, in just 15 epochs in deep residual U-net, the training accuracy is 98.89%. The recorded test accuracy is 91.29%.

The comparison table of performance of different methods is shown in Table I.

TABLE I. COMPARISION TABLE OF TEST ACCURACIES

Model	Dataset 1	Dataset 2
U-net [12]	0.8678	0.8960
Residual U-net [13]	0.8910	0.9129
Y-net [14]	0.8875	0.9057
Residual Y-net	**0.9039**	**0.9360**

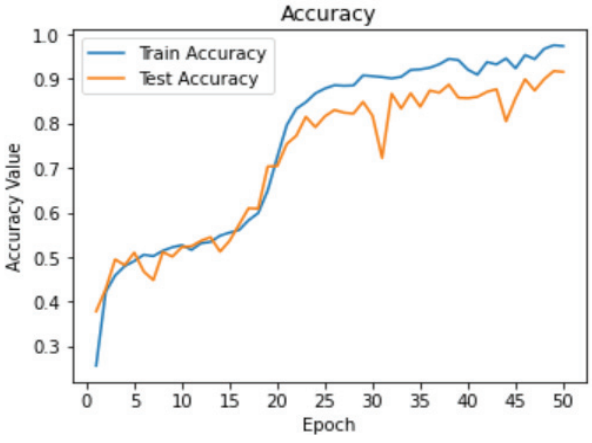


Fig. 5 Accuracy of Residual Y-net on Dataset 1

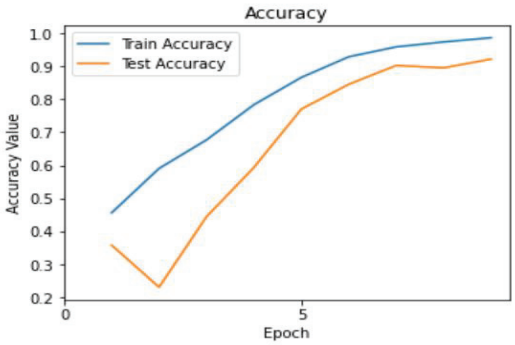


Fig. 6 Accuracy of Residual Y-net on Dataset 2

The number of epochs for training the model from U-net, Y-net, Residual U-net to Residual Y-net decreases significantly, while the test accuracy increases. Residual Units are proven better instead of traditional convolutional blocks in the networks. Two encoders instead of one encoder help in better feature mapping, and so it produces better segmentation and detection results. The model's performance got better due to the balanced medium size dataset.

V. CONCLUSION

The proposed method's sole aim is to maintain more features in a condensed form such that segmentation results for image classification can be enhanced when decoding in the decoder. Residual Y-net attained test accuracy of 93.60% in 9 epochs on a medium-size Kaggle Dataset. It does not let to degrade features between low and high levels. Because it uses residual units, the weights of input layers are preserved, and two encoders assisted in the storage of extra features components. There are no restrictions for spatial values. The architecture is flexible in placing any spatial size of residual units. Future work will include testing the model on a much larger dataset. There will be experiments on bringing more modifications in the architecture to increase the accuracy of classifications. Further research will include ways to tackle unbalanced dataset and give better accuracies.

REFERENCES

[1] W. Guo, M. Li, Y. Dong, H. Zhou, Z. Zhang, C. Tian, R. Qin, H. Wang, Y. Shen, K. Du and others, "Diabetes is a risk factor for the progression and prognosis of COVID-19," *Diabetes/metabolism research and reviews,* vol. 36, p. e3319, 2020.

[2] L. Roncon, M. Zuin, G. Rigatelli and G. Zuliani, "Diabetic patients with COVID-19 infection are at higher risk of ICU admission and poor short-term outcome," *Journal of Clinical Virology,* vol. 127, p. 104354, 2020.

[3] Z. Zheng, F. Peng, B. Xu, J. Zhao, H. Liu, J. Peng, Q. Li, C. Jiang, Y. Zhou, S. Liu and others, "Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis," *Journal of Infection,* 2020.

[4] S. Edagawa, F. Kobayashi, F. Kodama, M. Takada, Y. Itagaki, A. Kodate, K. Bando, K. Sakurai, A. Endo, H. Sageshima and others, "Epidemiological features after emergency declaration in Hokkaido and report of 15 cases of COVID-19 including 3 cases requiring mechanical ventilation," *Global Health & Medicine,* 2020.

[5] L. Orioli, M. P. Hermans, J.-P. Thissen, D. Maiter, B. Vandeleene and J.-C. Yombi, "COVID-19 in diabetic patients: Related risks and specifics of management," in *Annales d'endocrinologie*, 2020.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:49:20 UTC from IEEE Xplore. Restrictions apply.

[6] M. Sen, S. Lahane, T. P. Lahane, R. Parekh and S. G. Honavar, "Mucor in a viral land: a tale of two pathogens," *Indian journal of ophthalmology,* vol. 69, p. 244, 2021.

[7] S. Waizel-Haiat, J. A. Guerrero-Paz, L. Sanchez-Hurtado, S. Calleja-Alarcon and L. Romero-Gutierrez, "A case of fatal rhino-orbital mucormycosis associated with new onset diabetic ketoacidosis and COVID-19," *Cureus,* vol. 13, 2021.

[8] M. Sen, S. G. Honavar, N. Sharma and M. S. Sachdev, "COVID-19 and Eye: A Review of Ophthalmic Manifestations of COVID-19," *Indian Journal of Ophthalmology,* vol. 69, p. 488, 2021.

[9] R. Klein, B. E. K. Klein and S. E. Moss, "Visual Impairment in Diabetes," *Ophthalmology,* vol. 91, pp. 1-9, 1984.

[10] *A Textbook of Clinical Ophthalmology, Third Edition.,* vol. 16, 2003, p. 47.

[11] R. N. Frank, "Diabetic retinopathy," *Progress in Retinal and Eye Research,* vol. 14, pp. 361-392, 1995.

[12] O. Ronneberger, P. Fischer and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Cham, 2015.

[13] 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)*, 2016.

[14] S. Mehta, E. Mercan, J. Bartlett, D. L. Weaver, J. G. Elmore and L. G. Shapiro, "Y-Net: Joint Segmentation and Classification for Diagnosis of Breast Biopsy Images," *CoRR,* vol. abs/1806.01313, 2018.

[15] Y. Chen, X. Jin, B. Kang, J. Feng and S. Yan, "Sharing Residual Units Through Collective Tensor Factorization in Deep Neural Networks," *CoRR,* vol. abs/1703.02180, 2017.

[16] A. Samanta, A. Saha, S. C. Satapathy, S. L. Fernandes and Y.-D. Zhang, "Automated detection of diabetic retinopathy using convolutional neural networks on a small dataset," *Pattern Recognition Letters,* vol. 135, pp. 293-298, 2020.

[17] W. Chen, B. Yang, J. Li and J. Wang, "An Approach to Detecting Diabetic Retinopathy Based on Integrated Shallow Convolutional Neural Networks," *IEEE Access,* vol. 8, p. 178552–178562, 2020.

[18] G. Wang, J. Sun, J. Ma, K. Xu and J. Gu, "Sentiment classification: The contribution of ensemble learning," *Decision Support Systems,* vol. 57, pp. 77-93, 2014.

[19] B. Tymchenko, P. Marchenko and D. Spodarets, "Deep Learning Approach to Diabetic Retinopathy Detection," *CoRR,* vol. abs/2003.02261, 2020.

[20] R. Patel and A. Chaware, "Transfer Learning with Fine-Tuned MobileNetV2 for Diabetic Retinopathy," in *2020 International Conference for Emerging Technology (INCET)*, 2020.

[21] Z. Zeng, W. Xie, Y. Zhang and Y. Lu, "RIC-Unet: An improved neural network based on Unet for nuclei segmentation in histology images," *Ieee Access,* vol. 7, p. 21420–21428, 2019.

[22] X. Xiao, S. Lian, Z. Luo and S. Li, "Weighted res-unet for high-quality retina vessel segmentation," in *2018 9th international conference on information technology in medicine and education (ITME)*, 2018.

[23] J. Hu, H. Wang, S. Gao, M. Bao, T. Liu, Y. Wang and J. Zhang, "S-unet: A bridge-style u-net framework with a saliency mechanism for retinal vessel segmentation," *IEEE Access,* vol. 7, p. 174167–174177, 2019.

[24] S. Suri, A. Gupta and K. Sharma, "Comparative Analysis of Ranking Algorithms Used On Web," *Annals of Emerging Technologies in Computing (AETiC),* vol. 4, 2020.

[25] A. K. Tripathi, K. Sharma, M. Bala, A. Kumar, V. G. Menon and A. K. Bashir, "A Parallel Military-Dog-Based Algorithm for Clustering Big Data in Cognitive Industrial Internet of Things," *IEEE Transactions on Industrial Informatics,* vol. 17, pp. 2134-2142, 2021.

[26] S. Woo, J. Park, J.-Y. Lee and I. S. Kweon, "Cbam: Convolutional block attention module," in *Proceedings of the European conference on computer vision (ECCV)*, 2018.

[27] C. Guo, M. Szemenyei, Y. Yi, W. Wang, B. Chen and C. Fan, "Sa-unet: Spatial attention u-net for retinal vessel segmentation," in *2020 25th International Conference on Pattern Recognition (ICPR)*, 2021.

[28] Y. Zong, J. Chen, L. Yang, S. Tao, C. Aoma, J. Zhao and S. Wang, "U-net Based Method for Automatic Hard Exudates Segmentation in Fundus Images Using Inception Module and Residual Connection," *IEEE Access,* vol. 8, p. 167225–167235, 2020.

[29] C. Guo, M. Szemenyei, Y. Hu, W. Wang, W. Zhou and Y. Yi, "Channel Attention Residual U-Net for Retinal Vessel Segmentation," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021.

[30] A. Foo, W. Hsu, M. L. Lee, G. Lim and T. Y. Wong, "Multi-Task Learning for Diabetic Retinopathy Grading and Lesion Segmentation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020.

[31] J. E. McManigle, R. R. Bartz and L. Carin, "Y-Net for Chest X-Ray Preprocessing: Simultaneous Classification of Geometry and Segmentation of Annotations," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2020.

[32] A. Mohammed, S. Yildirim, I. Farup, M. Pedersen and Ø. Hovde, "Y-net: A deep convolutional neural network for polyp detection," *arXiv preprint arXiv:1806.01907,* 2018.

[33] Z. Zhang, Q. Liu and Y. Wang, "Road extraction by deep residual u-net," *IEEE Geoscience and Remote Sensing Letters,* vol. 15, p. 749–753, 2018.

[34] "APTOS 2019 Blindness Detection," 2019. [Online]. Available: https://www.kaggle.com/c/aptos2019-blindness-detection/. [Accessed 2019 Jun 10].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:49:20 UTC from IEEE Xplore. Restrictions apply.

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)
Pune, India. Oct 29-30, 2021

2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON) | 978-1-6654-2503-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/SMARTGENCON51891.2021.9645751

Sentiment Analysis on Twitter using Ordinal Regression

Moin Ahmed
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
moinahmed_2k18co209@dtu.ac.in

Mohit Goel
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
mohitgoel_2k18co206@dtu.ac.in

Raju Kumar
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
yadavraju03@yahoo.com

Dr. Aruna Bhat
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
aruna.bhat@dtu.ac.in

***Abstract*— With around 186 million daily active users, Twitter is a popular social media platform with increasing user growth every year. The top three twitter markets are in the US, Japan, and India. Users tweet on various topics ranging from sports, politics, finance to entertainment. Twitter is also used by businesses to advertise their products and services. These tweets contain a vast amount of information that can be used in many different ways, from targeting advertisements, understanding customer sentiment, and analyzing users' opinions about specific people or events. This research aims to apply sentiment analysis on Twitter data by Ordinal Regression. We use the *twitter_samples* from NLTK Corpus containing positive and negative tweets for our study. We use Random Forest (RF), Support Vector Machine (SVM), and Multinomial Logistic Regression (MLR) classifiers to classify the tweet sentiment into five categories. Results show that this approach outperformed the related work in terms of accuracy. This improvement was possible by applying TextBlob to find the polarity of the tweets and using lemmatization instead of stemming.**

Keywords*— *Sentiments, Twitter, Online Social Media, Machine Learning, Hate Speech

I. INTRODUCTION

Sentiment analysis, popularly called opinion mining, is a technique to gather personal information to help understand the sentiment from a textual data source. Brands and businesses often use it to understand the consumers and get feedback on their products to make better business decisions and enhance customer satisfaction. Researchers apply a Rule-based approach or an Automatic approach to perform such analysis. A rule-based approach uses Natural Language Processing (NLP) techniques and a set of rules created for the analysis. Automatic approaches, on the other hand, use Machine Learning techniques and don't use special rules. We will use the latter approach for studying our objective.

Twitter is a microblogging and social networking platform. With the increasing popularity and user engagement, Twitter has become a perfect platform where users express their opinions. Thus Twitter sentiment analysis is getting a lot of attention. Such a study can be beneficial because it collects and classifies public opinion by analyzing broad social data. Thus, it can be studied further for research in this area.

Twitter data, however, has some features that, in comparison to the analysis of other forms of data, create difficulties in performing sentiment analysis. There are a lot of challenges in this analysis. First, a tweet cannot have more than 140 characters in it. Second, most tweets are written in the informal form of a language (like English). These tweets also include slang words. Due to these problems, studies are being conducted on this topic.

This study mainly focuses on the polarity of the tweets and classifies the sentiment into five categories using Random Forest, Support Vector Machine, and Multinomial Logistic Regression machine learning algorithms.

The rest of the paper is as follows: Section 2 contains a review of related work done in this field and the various approaches by different researchers. Section 3 contains a detailed explanation of the methodology we have adopted for the sentiment analysis of Twitter data, divided into four parts: collecting data, preprocessing the data, extracting useful features, and finally classifying the data. Finally, section 4 includes the summary of the research paper and the conclusion we made from conducting the research.

II. RELATED WORK

A lot of work has been done in Sentiment Analysis of Twitter Data in recent years.

Twitter data is not directly usable in a study and needs to be preprocessed. Preprocessing includes removing URLs, numbers, stop words, and repeated letters, expanding acronyms. By preprocessing the tweets, both the accuracy and F1 measure increase [1].

Some researchers have also worked on Lexicon-based Sentiment Analysis, which uses tokenization, eliminating stop words and stemming. However, the accuracy is lower as compared to the work which used machine learning [2].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:17:28 UTC from IEEE Xplore. Restrictions apply.

Sentiment analysis using Machine Learning algorithms has been quite popular. M. Wongkar et al., 2019 [3] conducted sentiment analysis of tweets during the 2019 presidential elections held in the Republic of Indonesia. Naive Bayes Algorithm was used along with python programming language and achieved an accuracy of 80.90%. The study found a positive sentiment for the different pairs of candidates.

Md Rafiqul Islam et al., 2018 [4] investigated the problem of depression among Facebook users. They applied the KNN algorithm on Facebook Data to detect depressive emotions. Their work gave results between 60-70% on various matrices.

Megha Rathi et al. [5] used ensemble techniques for the sentiment analysis of tweets to improve the efficiency and reliability of their model. SVM algorithm performed the best with 82% accuracy.

Deep Learning has also proved useful in this area. A combination of CNN and a category of RNN, the LSTM network was used by Dionysis Goularas et al., 2018 [6] to study sentiment analysis. Deep Convolutional Neural Network [7], Back-propagation Neural Network [8], and some Hybrid Deep Learning Model [9] have also been researched.

Twitter sentiment analysis can be applied to a lot of applications like Sarcasm Detection [10],[21], Fake News Detection [11], Financial Sentiment Analysis [12], Cryptocurrency Price Forecasting [13], and Prediction of Election Results [14].

III. METHODOLOGY

This section deals with the proposed architecture and discusses the result in detail. The proposed framework comprises the following steps, as shown in Fig. 1.

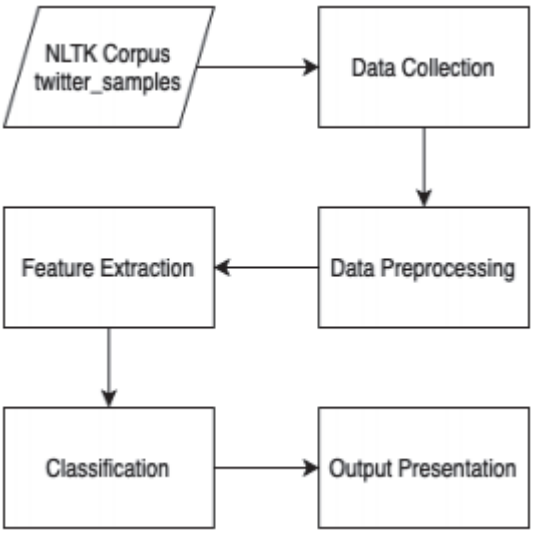


Fig. 1. Proposed Architecture

A. Data Collection

Data acquisition is the first step, in which we will collect tweets for sentiment analysis. We collected the data from NLTK Corpus twitter_samples, which contains 5000 positive and 5000 negative tweets [15].

B. Data Preprocessing

In the second step, the data will go through preprocessing to convert and clean the tweets to a dataset. But, first, we will separate the positive and negative tweets and process them separately to find the polarity distribution among positive and negative tweets and then follow the following process as shown in Fig. 2.

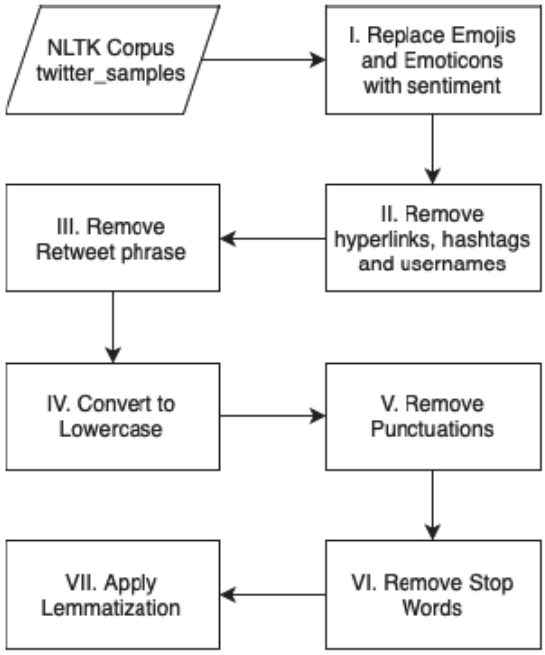


Fig. 2. Flowchart of Data Preprocessing

Firstly, we created a dictionary of emojis and their conveyed meaning. Then, the dictionary value replaced these emojis in the tweets. Emoticons were also replaced by the conveyed meaning using the emoji python package [16]. In stage II, we removed all the hyperlinks, hashtags, and usernames in the tweets. In stage III, the retweet phrases were removed. In stages IV and V, text was converted to lower case, and punctuations were removed. Stop words were removed from the text in stage VI.

At this stage, we have a lot of information in the data. Though present in different forms (synonyms or different word tenses), many words convey the same meaning or add the same information from a data point of view. We want to reduce these words down to their basic form, conveying enough info in a non-redundant manner. There are two ways to achieve the required reduction. Stemming is used to reduce words to their root (called stem) using suffix and prefix removal and does not consider the language aspect of the words. On the other hand, lemmatization reduces words that resemble the same meaning in a language to the root (called lemma) and works better. For example, *drink, drank, drunk and drinking* have the same lemma *drink*, but stemming will not do so.

In the last preprocessing stage, tweets are lemmatized instead of stemming from the inflected words' root form, using the WordNet module in NLTK library [17]. We used lemmatization instead of stemming because sometimes, the word loses its meaning in the stemming process. It may be faster than lemmatization, but we focused on accuracy.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:17:28 UTC from IEEE Xplore. Restrictions apply.

C. *Feature Extraction*

The third step includes getting the relevant features. We calculated the polarity of the tweets using TextBlob [18] python library. TextBlob is a library used for different processes of NLP. By applying TextBlob to the data, we got the polarity and sentiment of a tweet. Polarity ranges from -1 to 1. A -1 polarity represents a negative sentiment, while a +1 polarity represents a positive sentiment. We only used polarity as a feature for the study. After that, we balanced and scored the tweets.

The score of a tweet was calculated using polarity. A tweet with polarity above 0.5 was given a score of 2 and classified as a 'highly positive' tweet. A tweet with a polarity between 0 and 0.5 was scored with 1 and categorized as 'moderately positive.' A polarity equal to 0.0 would fetch a 0 score and the term 'neutral.' Similarly, -1 and -2 scores were given to tweets with the polarity between -0.5 and 0 or below -0.5 respectively under 'moderately negative' and 'highly negative' category.

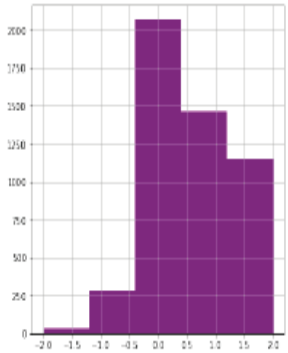


Fig. 3. Positive Tweets Score Distribution

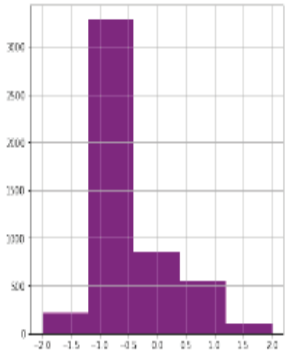


Fig.4. Negative Tweets Score Distribution

The scores of positive and negative tweets were plotted. We found that even though the tweets were labeled as positive and negative, they can still be divided into various categories using the score. Fig. 3. and 4. shows the distribution of score in positive and negative tweets.

D. *Classification and Result Analysis*

This subsection deals with different machine learning classifiers used in our framework and discusses their results. Different supervised learning classifiers are applied to preprocessed tweets and classify the tweets into five categories. These are highly positive, moderately positive, neutral, moderate negative, and high negative.

The dataset was formed by combining positive and negative tweets and the tweets' score as a single dataset. Table 1. shows the number of tweets per score in the final dataset. Then it was split into test and train datasets in a ratio of 7:3.

TABLE I. NUMBER OF TWEETS PER SCORE VALUE

Score	**-2**	**-1**	**0**	**1**	**2**
No. of Tweets	***252***	***3562***	***2916***	***2014***	***1256***

Our framework mainly focused on standard classifiers typically used in text classification like Random Forest(RF), Support Vector Machine(SVM), and Multinomial Logistic Regression(MLR)

Random Forest is a supervised machine learning algorithm that contains decision trees with different hyper-parameters and predicts majority votes.

The Support Vector Machine algorithm is a famous algorithm used for classification. Its objective is to locate a hyperplane in an N-dimensional space. The hyperplane has a maximum possible margin with the data points and helps in classification.

Unlike binary classification problems, where there are only two possible outcomes, there can be more than two possible outcomes in a multi-class classification problem. So, Multinomial Logistic Regression is similar to Logistic Regression and is used for multi-class classification problems.

We created these three classifiers and trained them on the dataset using Python programming language and Scikit-Learn [19] Machine Learning library.

The performance of these classifiers is evaluated using standard evaluation matrices like Accuracy, Precision, Recall, and F1- Score, as shown in Table II. The value of each evaluation matrices to the corresponding classifier is also mentioned in Table II and Fig. 5.

TABLE II. PERFORMACE OF OUR FRAMEWORK

Classifier	**Accuracy**	**Precision**	**Recall**	**F1-Score**
RF	*0.84*	*0.87*	*0.85*	*0.85*
SVM	*0.86*	*0.88*	*0.87*	*0.87*
MLR	*0.81*	*0.85*	*0.82*	*0.83*

Table II and Fig. 5 exhibit that the SVM classifier's overall performance is preeminent compared to other classifiers. SVM classifier achieves up to 86% accuracy and 88% precision and is nearly followed by the Random Forest and Multinomial logistic regression. Thus, the performance achieved by our approach is a promising outcome in this particular field.

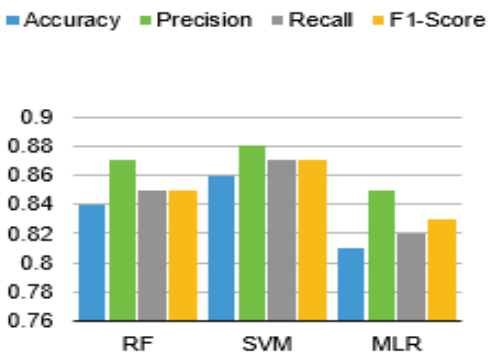


Fig. 5. Performance Analysis of Classifiers

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:17:28 UTC from IEEE Xplore. Restrictions apply.

In the last section, we compared our approach by S. Elbagir et al., 2019 [20], which used the same dataset; The Model-wise comparison of accuracy is shown in Table III. In the MLR classifier, there is an improvement of 14.60% accuracy and 4.65% improvement in the SVM classifier. This improvement was possible with the help of the TextBlob library and the lemmatization process.

TABLE III. ACCURACY COMPARISION WITH REF. [20]

Classifier	Our Approach	Ref. [20]
RF	*84.86%*	*83.20%*
SVM	*86.60%*	*81.95%*
MLR	*81.80%*	*67.20%*

IV. CONCLUSION

To summarize our work, we did sentiment analysis on Twitter data. The sentiment can be one of the five categories. We gathered data from *twitter_samples* NLTK Corpus, which contains positive and negative tweets. We preprocessed the data to get the relevant information only and used lemmatization. We got the polarity of the tweets using the TextBlob library. They were categorized into five categories. Finally, three ML algorithms were applied to the data.

SVM classifier performed the best in Twitter Sentiment Analysis with an accuracy of 86.60%, followed by Random Forest classifier (84.86%) and Multinomial Logistic Regression classifier (81.80%). While using SVM, we used GridSearchCV to find the best hyperparameters, resulting in improved accuracy of our Support Vector Classifier. We found that at 'C':1000 and 'gamma': 0.001, our model works best.

REFERENCES

[1] Z. Jianqiang and G. Xiaolin, "Comparison Research on Text Pre-processing Methods on Twitter Sentiment Analysis," in IEEE Access, vol. 5, pp. 2870-2879, 2017, doi: 10.1109/ACCESS.2017.2672677.

[2] Kusrini and M. Mashuri, "Sentiment Analysis In Twitter Using Lexicon Based and Polarity Multiplication," 2019 International Conference of Artificial Intelligence and Information Technology (ICAIIT), 2019, pp. 365-368, doi: 10.1109/ICAIIT.2019.8834477.

[3] M. Wongkar and A. Angdresey, "Sentiment Analysis Using Naive Bayes Algorithm Of The Data Crawler: Twitter," 2019 Fourth International Conference on Informatics and Computing (ICIC), 2019, pp. 1-5, doi: 10.1109/ICIC47613.2019.8985884.

[4] M. R. Islam, A. R. M. Kamal, N. Sultana, R. Islam, M. A. Moni and A. ulhaq, "Detecting Depression Using K-Nearest Neighbors (KNN) Classification Technique," 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2), 2018, pp. 1-4, doi: 10.1109/IC4ME2.2018.8465641

[5] M. Rathi, A. Malik, D. Varshney, R. Sharma and S. Mendiratta, "Sentiment Analysis of Tweets Using Machine Learning Approach," 2018 Eleventh International Conference on Contemporary Computing (IC3), 2018, pp. 1-3, doi: 10.1109/IC3.2018.8530517.

[6] D. Goularas and S. Kamis, "Evaluation of Deep Learning Techniques in Sentiment Analysis from Twitter Data," 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML), 2019, pp. 12-17, doi: 10.1109/Deep-ML.2019.00011.

[7] Z. Jianqiang, G. Xiaolin and Z. Xuejun, "Deep Convolution Neural Networks for Twitter Sentiment Analysis," in IEEE Access, vol. 6, pp. 23253-23260, 2018, doi: 10.1109/ACCESS.2017.2776930.

[8] N. A. Setyadi, M. Nasrun and C. Setianingsih, "Text Analysis For Hate Speech Detection Using Backpropagation Neural Network," 2018 International Conference on Control, Electronics, Renewable Energy and Communications (ICCEREC), 2018, pp. 159-165, doi: 10.1109/ICCEREC.2018.8712109.

[9] M. U. Salur and I. Aydin, "A Novel Hybrid Deep Learning Model for Sentiment Classification," in IEEE Access, vol. 8, pp. 58080-58093, 2020, doi: 10.1109/ACCESS.2020.2982538.

[10] L. H. Son, A. Kumar, S. R. Sangwan, A. Arora, A. Nayyar and M. Abdel-Basset, "Sarcasm Detection Using Soft Attention-Based Bidirectional Long Short-Term Memory Model With Convolution Network," in IEEE Access, vol. 7, pp. 23319-23328, 2019, doi: 10.1109/ACCESS.2019.2899260.

[11] O. Ajao, D. Bhowmik and S. Zargari, "Sentiment Aware Fake News Detection on Online Social Networks," ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019, pp. 2507-2511, doi: 10.1109/ICASSP.2019.8683170.

[12] S. Sohangir, N. Petty and D. Wang, "Financial Sentiment Lexicon Analysis," 2018 IEEE 12th International Conference on Semantic Computing (ICSC), 2018, pp. 286-289, doi: 10.1109/ICSC.2018.00052.

[13] A. Jain, S. Tripathi, H. D. Dwivedi and P. Saxena, "Forecasting Price of Cryptocurrencies Using Tweets Sentiment Analysis," 2018 Eleventh International Conference on Contemporary Computing (IC3), 2018, pp. 1-7, doi: 10.1109/IC3.2018.8530659.

[14] M. Bilal, A. Gani, M. Marjani and N. Malik, "Predicting Elections: Social Media Data and Techniques," 2019 International Conference on Engineering and Emerging Technologies (ICEET), 2019, pp. 1-6, doi: 10.1109/CEET1.2019.8711854.

[15] "Twitter HOWTO", nltk.org, 2021, [Online], Available: http://www.nltk.org/howto/twitter.html. Accessed 4 Jun 2021.

[16] Kim, Taehoon. "emoji 1.2.0", pypi, 27 Jan. 2021, [Online], Available: https://pypi.org/project/emoji/

[17] "nltk.stem package", nltk.org, 2021, [Online], Available:https://www.nltk.org/api/nltk.stem.html#nltk.stem.wordnet.WordNetLemmatizer. Accessed 4 Jun 2021.

[18] Loria, Steven. "TextBlob: Simplified Text Processing", textblob.readthedocs.io, 2021, [Online], Available: https://textblob.readthedocs.io/en/dev/. Accessed 4 Jun. 2021,

[19] "scikit-learn Machine Learning in Python", scikit-learn, 2021, [Online], Available: https://scikit-learn.org/stable/#. Accessed 4 Jun. 2021

[20] S. E. Saad and J. Yang, "Twitter Sentiment Analysis Based on Ordinal Regression," in IEEE Access, vol. 7, pp. 163677-163685, 2019, doi: 10.1109/ACCESS.2019.2952127.

[21] R. Kumar and A. Bhat, "An Analysis On Sarcasm Detection Over Twitter During COVID-19," 2021 2nd International Conference for Emerging Technology (INCET), 2021, pp. 1-6, doi: 10.1109/INCET51464.2021.9456392.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 28,2021 at 06:17:28 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/TMAG.2021.3131511, IEEE Transactions on Magnetics

Significant Improvement in Magnetic and Magnetoelectric Characteristics of (0.95-x)$Ba_{0.9}Ca_{0.1}TiO_3$-0.05$Na_{0.5}Bi_{0.5}TiO_3$-x$CoFe_2O_4$ ceramics

Aditya Jain[1], *Member, IEEE*, Ajay Kumar[2], *Senior Member, IEEE* and Neha Gupta[3], *Member, IEEE*

[1] Electronics and Telecommunication Engineering Department, Symbiosis Institute of Technology, Pune, India
[2] Electronics and Communication Engineering Department, Jaypee Institute of Information Technology, Noida, India
[3]Applied Science and Humanity Department, ADGITM, New Delhi, India

Multi-functional materials with high magnetic, electrical, and magnetoelectric characteristics play an important role in development of data storage devices. Herein, a multi-phase (0.95-x)$Ba_{0.9}Ca_{0.1}TiO_3$-0.05$Na_{0.5}Bi_{0.5}TiO_3$-$x$$CoFe_2O_4$ ceramics were prepared by two-stage mechanochemical activation technique. The objective of this study is to develop a multifunctional material with improved multiferroic characteristics. The structural analysis by Rietveld method shows the presence of heterogeneous structure without any impurity phase. The incorporation of $CoFe_2O_4$ (CFO) in 0.95$Ba_{0.9}Ca_{0.1}TiO_3$-0.05$Na_{0.5}Bi_{0.5}TiO_3$ (BCT-NBT) has resulted in significant improvement in magnetic and magnetoelectric characteristics without significantly hampering the electrical properties. The sample with x = 0.11 has shown high saturation magnetization of 8.26 emu/g along with moderately high magnetoelectric coefficient of 5.411 mV/cm.Oe. In addition to this, the energy storage performance of all the studied samples have also been investigated and the optimized sample x = 0.11 presents a large discharge energy density of 2.249 J/cm^3 and moderately high energy efficiency of 63.8% at 388 kV/cm. The results obtained in the present work suggest a significant influence of CFO substitution on the magnetic characteristics of BCT-NBT ceramics, which subsequently indicate their potential applications in spintronics, magnetic memory storage, and electrical energy storage devices.

***Index Terms*—** Data storage, Dielectric, Ferroelectric, Magnetic, Magnetoelectric**.**

I. Introduction

Multiferroic magnetoelectric (ME) ceramics have attracted a widespread attention of researchers across the globe due to their potential applicability in the design and development of multifunctional devices [1, 2]. Number of technological applications are ever-growing and these applications require novel materials with enhanced properties. Further, another constraint for the development of efficient material is the cost effectiveness, which require the materials to be inexpensive. At present, multiferroic materials are considered as one of the most desirable candidates for the next generation electronic devices. Now a days, multiferroic ceramic materials are widely used in various commercial nanoelectronics based electronic circuits, such as magnetic field sensing, read heads, energy harvesting, sensors, non-volatile memories, high frequency induction, and spintronics [3, 4]. Multiferroic ME materials have the ability to switch the magnetic moment on the application of applied electric field and vice-versa. However, due to the chemical incompatibility of electrical and magnetic phases, the interplay between ferroelectricity and magnetism has remained an area of research in the past decade. A comprehensive literature survey reveals that there are very few single phase magnetoelectric materials and therefore, combination of magnetic, ferroelectric, and piezoelectric can be prepared to achieve a significant magnetoelectric effect. The ME coupling in heterostructure ceramics derives from the magnetic-electric-mechanical energy transfer via the interface among the magnetostrictive, ferroelectric, and piezoelectric systems [5]. Over the past few years, the realization of multi-property coupling in a material has led to various potential applications, which is primarily ascribed to development in the area of multifunctional materials and integrated devices.

Among the various available dielectric/ferroelectric materials, $BaTiO_3$ (BT) is one of the most versatile material due to its high dielectric constant, large saturation polarization, moderately high breakdown strength, and moderately large piezoelectric coefficient [6, 7]. Further, it has been observed that the electrical and magnetic characteristics of BT can be enhanced by substituting/doping it with appropriate compound [8-10]. There are several reports which suggests that a suitable amount of Ca substitution in BT lattice may result in significant improvement in chemical stability and electrical properties in comparison to pure BT [11, 12]. Therefore, in the present work, an optimized Ca modified BT composition (i.e., $Ba_{0.9}Ca_{0.1}TiO_3$) has been considered as the parent phase of synthesized ceramic.

$Na_{0.5}Bi_{0.5}TiO_3$ (NBT) is another well-known lead-free ferroelectric/piezoelectric ceramic with excellent electrical properties. NBT exhibit high saturation polarization and piezoelectricity compared to BT and therefore, an appropriate amount of NBT substation in BT can improve the energy storage performance and electromechanical characteristics of the material. A comprehensive literature survey indicates that there are various reports, which shows that a combination of BT and NBT may result in enhancement in energy storage performance and electromechanically properties in comparison to pure BT alone [13-16]. Keeping the aspects in mind, a

The work of Aditya Jain, Ajay Kumar, and Neha Gupta is supported by Symbiosis Institute of Technology, Pune, Jaypee Institute of Information Technology, Noida, and ADGITM, New Delhi.

The authors are with the Electronics and Telecommunication Engineering Department, Symbiosis Institute of Technology, Pune, India, Electronics and Communication Engineering Department, Jaypee Institute of Information Technology, Noida, India, and Applied Science and Humanity Department ADGITM, New Delhi, India. (e-mail: adi.28.jain@gmail.com; ajay.kumar@jiit.ac.in; gupta.neha@niecdelhi.ac.in).

0018-9464 (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 December 06,2021 at 06:37:35 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/TMAG.2021.3131511, IEEE Transactions on Magnetics

suitable amount of NBT has been combined with BCT to obtain a material with excellent dielectric, ferroelectric, and piezoelectric properties.

Ferrites ($CoFe_2O_4$, $NiFe_2O_4$, $MnFe_2O_4$, etc.) are known to exhibit excellent magnetic properties. In the past few decades, $CoFe_2O_4$ (CFO) has been extensively studied due to their outstanding magneto-crystalline anisotropy, mechanical hardness, excellent chemical stability, high saturation magnetization, and high coercive field [17, 18]. Among the various available magnetic oxides, CFO with inverse spinal crystal structure is one of the most commercialized magnetic ferrite materials. The magnetic anisotropy characteristic of CFO is significantly susceptible to synthesis conditions, oxidation states, Co/Fe/O ratio, and occupancy of cations at specific sites [19, 20]. CFO exhibit an inverse spinal crystal structure with *Fd3m* space group. The unit cell of CFO consists of cubic closed pack lattice, in which metal cations are arranged octahedrally or tetragonally by oxygen atoms [21]. In addition to this, CFO also exhibit high catalytic efficiency, high reaction activity, and large surface to volume ratio, which makes it an ideal candidate for developing multiferroic ceramics [22]. The combination of BT and CFO is known to induce ME characteristics, which is due to spinodal decomposition that prevents reaction between the constituent phases during the final heat treatment process [23, 24]. Ceramics consisting of BT and CFO have been widely utilized in numerous applications, such as spintronics, memories, sensors, electromagnetic interference filters, transducers, etc.

In the recent past, various BT based composites and solid solutions were prepared to improve the magnetic and multiferroic characteristics. Choudhari et al. investigated the $(1-x)Mn_{0.7}Zn_{0.3}Fe_2O_4+(x)BaTiO_3$ composites via sol gel auto ignition method and the optimized composition with $x = 0.75$ exhibit the magnetoelectric coefficient of 20.45 mV/cm·Oe [25]. In another report, Choudhari et al. studied the Strain mediated magnetoelectric properties of $(1-x)CoCr_{0.3}Fe_{1.7}O_4+(x)BaTiO_3$ ceramics and obtained a significant improvement in magnetoelectric coefficient (24.7 mV/cm·Oe) for the composition with $x = 0.40$ $BaTiO_3$ concentration [26]. Wadgane et al. reported the magnetoelectric characteristics of $(x)YFeO_3–(1-x)Y_3Fe_5O_{12}$ nanocomposites synthesized via sol–gel auto combustion route. The study revealed that at higher $Y_3Fe_5O_{12}$ concentration, the ferroelectric loops have become unsaturated and the maximum magnetoelectric coefficient were found to be 10.711 mV/cm·Oe for $x = 0.50$ [27]. All these reports indicate that a significant improvement in magnetoelectric coefficient can be achieved by suitable choice of substituent/dopant in an appropriate concentration.

In this investigation, a multi-phase $(0.95-x)Ba_{0.9}Ca_{0.1}TiO_3$-$0.05Na_{0.5}Bi_{0.5}TiO_3$-$xCoFe_2O_4$ ($x = 0.0$, 0.03, 0.07, 0.11, and 0.15) ceramics have been prepared by two-stage mechanochemical activation technique. All the three phases were individually prepared and subsequently combined and sintered at an optimized temperature to pure phase ceramics. Following the synthesis, the magnetic, ferroelectric, dielectric, electromechanical, and magnetoelectric properties were systematically studied. A few strategies have been used to attain ceramic consisting of BCT, NBT, CFO phases. The combination of these three versatile compounds has resulted in an excellent multiferroic ceramic with improved magnetic and magnetoelectric characteristics. In the present composition, particles of BCT-NBT acts as the seeds for the CFO particles, which allows a homogeneous dispersion of parent phase with the guest phase. An excellent miscibility among the three phases has resulted in a tight integration among the particles and emergence of magnetic coupling in the material. It has been reported that a significant improvement in ME coefficient can be achieved by systematically optimizing the synthesis conditions [28, 29].

II. Experimental Details

A. Synthesis process

A series of $(0.95-x)Ba_{0.9}Ca_{0.1}TiO_3$-$0.05Na_{0.5}Bi_{0.5}TiO_3$-$xCoFe_2O_4$ ($x = 0.0$, 0.03, 0.07, 0.11, and 0.15) ceramics were synthesized by a two stage mechanochemical activation method. Heat treatment temperature and duration have been optimized to obtain pure phase ceramics. All the raw materials were of analytical reagent (AR) grade purity and used after heating them at 100 ºC for 15 minutes. At first, the raw materials $BaCO_3$ (99.95%), $CaCO_3$ (99.99%), and TiO_2 (99.9%) were used to prepare BCT phase. The oxides were mixed in stoichiometric amount and ball milled at 450 RPM for 15 effective hours. An optimized ball to powder ration of 11:1 has been kept for milling the materials. A few drops of ethyl alcohol were mixed with the power raw materials to reduce the heat generation and friction inside the ball milling jar. The obtained mixture was dried and subsequently calcined at 1075 ºC for 3 hours to get the pure perovskite phase. Next, the powder is again grinded in a pestle mortar and sintering is carried out at 1240 ºC for 4 hours to finally obtain the BCT phase.

Next, $Na_{0.5}Bi_{0.5}TiO_3$ (NBT) has been synthesized by taking Na_2CO_3 (99.99%), Bi_2O_3 (99.95%), and TiO_2 (99.9%) in the required stoichiometric amount. The raw materials were ball milled followed by drying and calcination at 825 ºC for 2 hours. Afterwards, the material was granulated and sintered at 1075 ºC for 2 hours to obtain the pure phase NBT solid solution.

At last, $CoFe_2O_4$ (CFO) is prepared by taking CoO (99.99%), and Fe_2O_3 (99.95%) in the required amount. The raw materials were milled, which is followed by drying at 100 ºC for 10 minutes. Subsequently, the mixture was calcined at 1050 ºC for 6 hours. The calcined powder is granulated and sintered at 1225 ºC for 12 hours. At this instant, we have obtained a pure phase CFO. After the preparation of all the three phases, the materials were mixed in pre-decided amount and ball milled for effective 5 hours. The obtained powder material is combined with an appropriate amount of polyvinyl alcohol solution (6 wt.%) for preparing the pellet samples. Afterwards, the material is annealed at an optimized temperature of 675 ºC for 4 hours to obtain a ceramic without any impurity phase. After this, the samples were assigned the nomenclature as BNC0, BNC3, BNC7, BNC11, and BNC15 for $x = 0.0$, 0.03, 0.07, 0.11 and 0.15, respectively.

B. Characterization

The structural characterization of all the samples were carried out by XRD-7000 (Shimadzu Kyoto, Japan) X-ray diffractometer. JEOL JSM-7500F FESEM was used to analyze

0018-9464 (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 December 06,2021 at 06:37:35 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/TMAG.2021.3131511, IEEE Transactions on Magnetics

the microstructural properties of prepared ceramics. The particle morphology of optimized sample has been carried out by Talos F200s tunneling electron microscope. PPMS-9 vibrating sample magnetometer was used to study the B-H hysteresis loops of CFO substituted samples. QUINTESS ferroelectric test system was used to analyze the room temperature ferroelectric characteristics and breakdown strength of all the five synthesized samples. Temperature dependent dielectric investigation of the samples has been performed by keysight technologies E4990A impedance analyzer. Radiant Precision Premier II test system has been utilized to measure the room temperature bipolar strain curves of the prepared samples. Radiant's D_{33} ceramic piezoelectric test system has been used to measure the piezoelectric coefficient of the samples. Prior to piezoelectric measurement, all the samples were electrically poled at an optimized electric field by Marine India make electrical poling system. At last, the temperature dependent ME coefficient of CFO substituted samples has been measured by Quantum Design Super-ME-II magnetoelectric test system.

III. Results and Discussion

A. Phase and structural characteristic analysis

The phase and crystalline properties of all the samples were studied by analyzing the X-ray diffraction patterns of the prepared samples. Fig. S1 in supplementary data shows the XRD patterns of the five synthesized samples. It has been observed that small amount of NBT substitution did not affect the crystallinity of pure BCT, which is very much desirable for achieving improved characteristics. With increase in CFO substitution, few extra peaks started to appear, which solely belongs to CFO. The reference XRD peak patterns of BT, NBT, and CFO are shown for better understand the emergence of various peaks in the prepared samples. The Rietveld analysis of all the synthesized sample has been performed by *FullProf* software. The Rietveld analysis of all the five synthesized samples is shown in Fig. 1(a-e). For the Rietveld analysis, suitable phases of BT (JCPDS no.: 01-074-1956), NBT (JCPDS no.: 01-085-0530), and CFO (JCPDS no.: 00-003-0864) were considered as the reference diffraction pattern, and the investigation shows that all the samples are in pure phase. After the successful Rietveld analysis, it has been observed that the BCT and CFO phases belongs to *P*4*mm* and *Fd*3*m* space group, respectively. The crystalline properties of all the five synthesized samples are shown in Table 1. The Rietveld study of CFO substituted samples indicate that the materials are heterogeneous in phase and consists of tetragonal, rhombohedral, and cubic symmetry belonging to BCT, NBT, and CFO phases, respectively. Previous reports on multiferroic composites and solid solutions also indicate that followed approach can results in improvement in magnetic and magnetoelectric characteristics [30-32]. A small goodness of fit (χ^2) parameter for all the synthesized samples confirms the validity of proposed approach. In addition to this, the absence of any undesirable diffraction peak in all the CFO substituted samples indicate a good solubility as well as compatibility of the three phases. Further, it can be observed that the NBT peaks were present in BNC3 sample; however, with increase in CFO concentration, the peak has started to diminish, which may be attributed to decomposition of NBT phase in BCT. Fig. 1(f) shows the variation of lattice parameters ('*a*' and '*c*') with increasing CFO concentration. Interestingly, with increase in CFO concentration, there is an increase in lattice parameter '*c*' and reduction in lattice parameter '*a*', which indicate an increase in *c/a* ratio of BCT phase. The *c/a* ratio is a measure of tetragonality in BT based compounds, and a higher value indicate higher tetragonal distortion, which is very much necessary for maintaining high value of dielectric constant and saturation polarization in the material. The results obtained in the present work are in agreement with previous reports on BT based ceramics [33, 34]. The change in tetragonality of the materials is also visible in the raw XRD patterns (Fig. S1 in supplementary data) of the samples, where an increase in peak splitting can be observed with increasing CFO substitution.

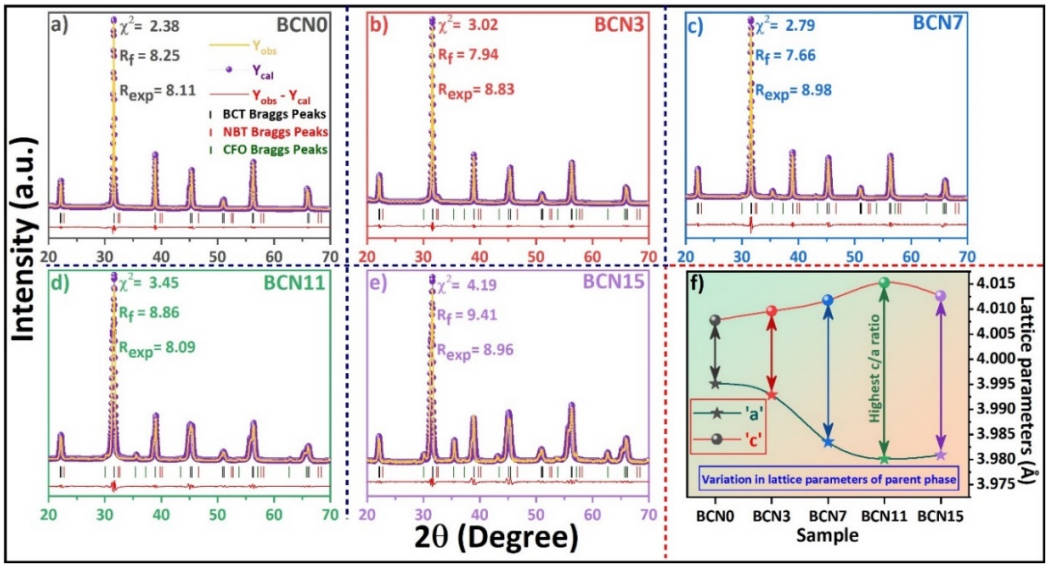


Fig. 1. Rietveld analysed XRD patterns of (0.95-*x*)BCT-0.05NBT-*x*CFO ceramic samples: (a) BNC0, (b) BNC3, (c) BNC7, (d) BNC11, and (e) BNC15. (f) Variation in the lattice parameters ('a' and 'c') of parent phase (i.e., BCT).

0018-9464 (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 December 06,2021 at 06:37:35 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/TMAG.2021.3131511, IEEE Transactions on Magnetics

Table 1.Space group, crystal symmetry, lattice parameters, unit cell volume, Rietveld analyzed parameters, material's theoretical density, experimental density, and weight fraction for the (0.95-*x*)BCT-0.05NBT-*x*CFO ceramic samples.

Sample		Space group	Crystal structure	Lattice parameters (Å)	Cell volume (Å^3)	χ^2	R_{exp}	R_f	R_{Bragg}	Theoretical density (gm/cm^3)	Experimental density (gm/cm^3)	Weight fraction (%)
BNC0		*P*4*mm*	Tetragonal	a = 3.9951 c = 4.0078	63.9870	2.38	8.11	8.25	2.97	5.77	5.49	93.07
		*R*3*c*	Rhombohedral	a = 5.4829 c = 6.7611	203.253				4.08			4.57
BNC3	BCT phase	*P*4*mm*	Tetragonal	a = 3.9929 c = 4.0096	63.9260	3.02	8.83	7.94	2.74	5.39	5.21	90.03
	NBT phase	*R*3*c*	Rhombohedral	a = 5.4785 c = 6.7595	202.879				4.48			4.48
	CFO phase	*Fd*3*m*	Cubic	a = 8.3729	586.985				3.96			2.89
BNC7	BCT phase	*P*4*mm*	Tetragonal	a = 3.9835 c = 4.0118	63.6603	2.79	8.98	7.66	3.05	5.23	5.15	86.17
	NBT phase	*R*3*c*	Rhombohedral	a = 5.4795 c = 6.7562	202.935				3.78			4.18
	CFO phase	*Fd*3*m*	Cubic	a = 8.3655	585.430				4.41			6.68
BNC11	BCT phase	*P*4*mm*	Tetragonal	a = 3.9801 c = 4.0153	63.6071	3.45	8.09	8.86	3.78	5.58	5.41	82.41
	NBT phase	*R*3*c*	Rhombohedral	a = 5.4766 c = 6.7589	202.720				3.91			4.15
	CFO phase	*Fd*3*m*	Cubic	a = 8.3621	584.717				4.21			10.52
BNC15	BCT phase	*P*4*mm*	Tetragonal	a = 3.9809 c = 4.0126	63.5899	4.19	8.96	9.41	4.35	5.76	5.59	78.66
	NBT phase	*R*3*c*	Rhombohedral	a = 5.4739 c = 6.7605	202.568				4.87			4.09
	CFO phase	*Fd*3*m*	Cubic	a = 8.3605	584.381				4.19			14.39

B. Microstructure investigation

I. FESEM study

A detailed microstructural study of prepared ceramics have been performed by examining the FESEM images. The FESEM images of bare BCT along with five substituted samples are shown in Fig. 2(a-f). The bare BCT ceramic exhibit slightly porous morphology along with small grain size in comparison to other substituted compositions in the work. The NBT and CFO substitution in BCT has a significant effect on the microstructural characteristics and material density in comparison to bare BCT. With increase in CFO substitution, the grain size as well as the material density first decrease and then increases on increasing the CFO substitution. It can be observed that all the CFO substituted samples are relatively dense and there are very few pores. The distribution of grains in the observed area has been calculated by *Imagej* software and depicted in Fig. 2(g-l). The samples BNC0, BNC3, BNC7, BNC11, and BNC15 show an average grain size of 693, 609, 1142, 1077, and 1198 nm, respectively with an accuracy of ±10 nm. One of the reasons for increase in grain size is the increase in number of oxygen vacancies with CFO substitution. The increase in grain size with increase in CFO concentration indicates that CFO acts as grain growth promoter. Further, the changes in concentration of metal oxides and defect chemistry of CFO substituted material have led to reduction in mass transportation and subsequent increase in grain size of the material [35]. In order to obtain the best of electrical and magnetic properties of BCT and CFO phase, respectively, an additional annealing is carried out at 675 ºC for 4 hours, which may have resulted in the emergence of liquid phase at grain boundaries leading to reduction in ion diffusion and consequently higher grain growth. The theoretical density of all the samples has been evaluated from the Rietveld study. Further, the Archimedes' technique of density measurement was employed to measure the experimental material density of the samples. It should be noted that both theoretical density and experimental density are in agreement with each other, which again verify the validity of Rietveld analysis. The addition of relatively conductive CFO phase is likely to decrease the breakdown strength of BNC*x* ceramics; however, the presence of dense microstructure in higher concentration CFO ceramics will help in moderating the breakdown strength.

II. TEM, HRETM, and SAED investigation

The particle morphology and structural characteristics of optimized BNC11 ceramic has been analyzed by TEM, HRTEM, and SAED analysis. Fig. 3(a) shows the TEM image of BNC11 sample. It should be noted that we have taken several TEM images and obtained almost identical results each time. It can be observed from the TEM image that BNC11 consists of three types of particles: light, moderately dark, and highly dark. The majority of BNC11 sample contains light particles and are distributed evenly throughout the observed region. The HRTEM image of a light particle (Fig. 3(b)) indicate collaterally situated lattice fringes, which suggests the high crystallinity. The distance between the collateral fringes were found to be ~1.41 Å, which corresponds to [2 2 0] lattice plane of BCT phase. This analysis has been done by comparing the obtained results with same benchmark pattern which was used for Rietveld analysis. Next, various moderately dark particles were analyzed by HRTEM and one of them is shown in Fig.

0018-9464 (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 December 06,2021 at 06:37:35 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/TMAG.2021.3131511, IEEE Transactions on Magnetics

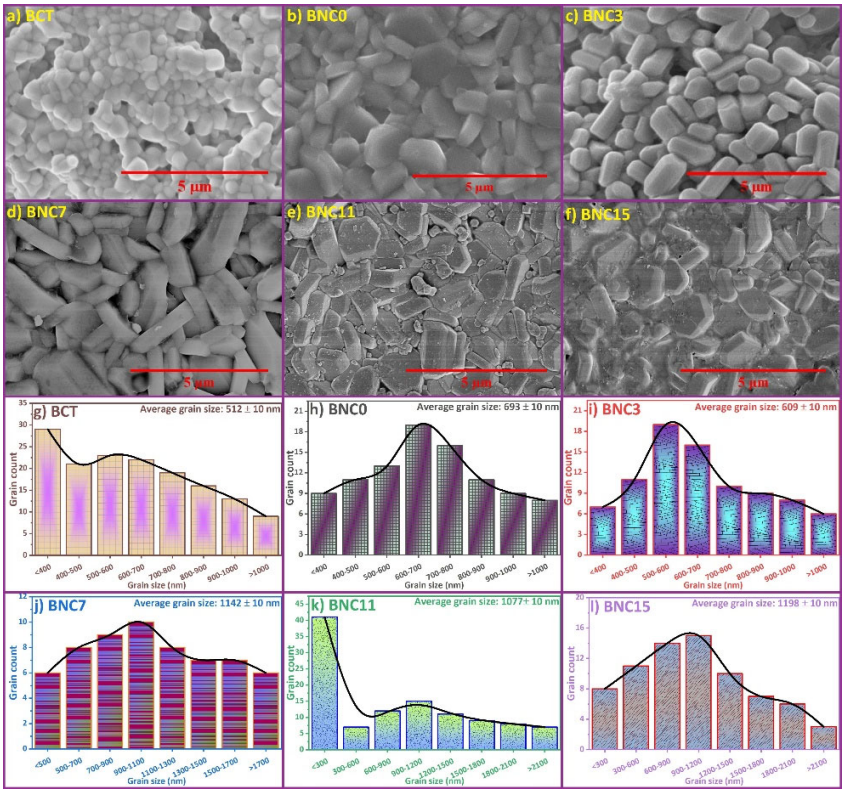


Fig. 2. Cross-sectional FESEM micrograph and corresponding grain size distribution of (0.95-*x*)BCT-0.05NBT-*x*CFO ceramic samples: (a and g) Pure BCT, (b and h) BNC0, (c and i) BNC3, (d and j) BNC7, (e and k) BNC11, and (f and l) BNC15.

3(c). Again, a suitable comparison has been made with the NBT phase and it is found that the fringes with an inter-fringe distance of ~1.94 Å with [2 0 2] lattice plane belongs to NBT phase. At last, highly dark particles were examined via HRTEM (Fig. 3(d)) and it has been observed that the particle contains collaterally aligned lattice fringes in addition to twinning of lattice fringes at some places. The lattice fringes with an inter-fringe distance of 1.48 Å are associated with [4 4 0] lattice plane of CFO phase. It is worth noting that the analysis has been carried out in various highly dark particles and an almost identical results have been obtained. The obtained results from TEM and HRTEM have been further verified by SAED analysis. Fig. 3(e) shows the SAED pattern of an area where light and highly dark particle infuse with each other. The presence of clearly visible bright rings confirms the polycrystalline nature of BNC11 ceramic. The presence of

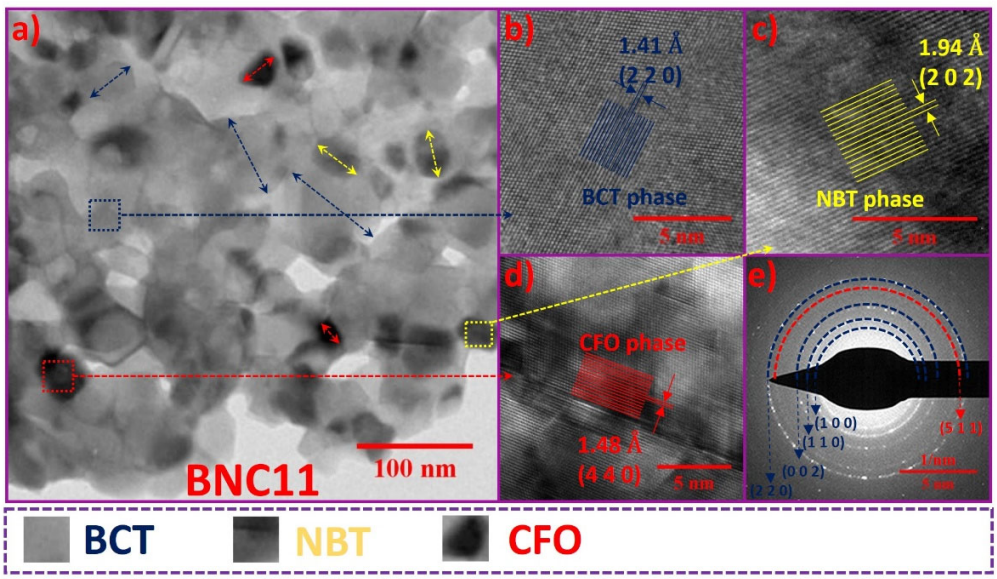


Fig. 3. TEM, HRETM, and SAED micrographs of BNC11 sample: (a) TEM, (b, c, and d) HRTEM, and (e) SAED.

0018-9464 (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 December 06,2021 at 06:37:35 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/TMAG.2021.3131511, IEEE Transactions on Magnetics

circular rings in the SAED pattern indicates the polycrystalline nature of the material. The bright rings in the SAED pattern were marked with appropriate lattice plane by measuring the d spacing. The blue colored rings in the SAED pattern indicate the presence of BCT phase and red colored ring is associated with CFO phase. The HRTEM and SAED analysis confirm the presence of multiple phases in the BNC11 ceramic, which will be beneficial in obtaining improved multiferroic characteristics.

C. *Magnetic properties*

Fig. 4(a-d) depict the room temperature magnetic hysteresis loops (M-H) of all four CFO substituted samples. The M-H loop of BNC0 sample is not shown, as both BCT and NBT do not have any significant magnetic characteristics. Even a small amount of CFO substitution has resulted in emergence of ferromagnetic phenomena in the ceramics. The M-H loops were measured at a maximum applied magnetic field of 20 kOe. With increase in CFO substitution, the loops have become more and more saturated, which may be ascribed to feeble intrinsic diamagnetic behavior of BCT phase and exchange energy interaction at higher applied magnetic field. The primary source of ferromagnetism in BT based multiferroic ceramics is the F-centre exchange phenomena, which is dependent on vacancies induced by lattices (Primarily oxygen vacancies) and lattice defects [36, 37]. In F-centre exchange phenomena, two metal ions (e.g. Co^{2+}) interact with each other via available oxygen vacancies and results in polaronic effect [38, 39]. During this process, an electron is entrapped in between the oxygen vacancies, that give rise to the occurrence of F-centre exchange phenomena, which is nearly identical to magnetic bound polarons. Therefore, it can be emphasized that D-centre bound polarons are responsible for the occurrence of ferromagnetism in CFO substituted ceramics. The saturation magnetization (M_s), remnant magnetization (M_r), and coercive magnetic field

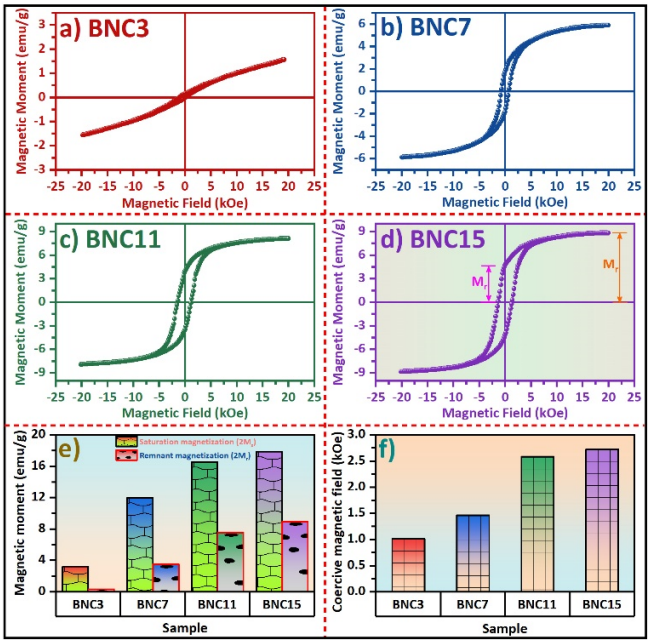


Fig. 4. B-H loops for the CFO substituted samples: (a) BNC3, (b) BNC7, (c) BNC11, and (d) BNC15. (e and f) Schematics show the comparison of saturation magnetization, remnant magnetization, and coercive magnetic field.

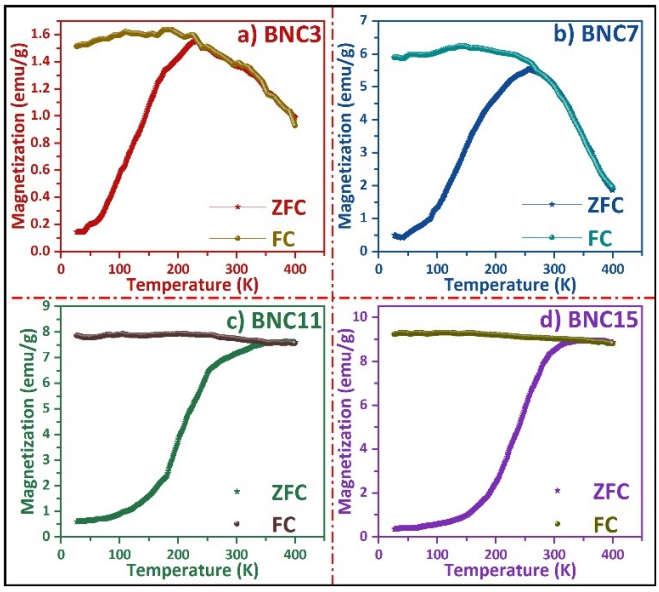


Fig. 5. M-T curves carried out with zero-field-cooling (ZFC) and field-cooling (FC) (a) BNC3, (b) BNC7, (c) BNC11, and (d) BNC15.

(H_c) increases significantly with increase in CFO substitution. Fig. 4(e-f) shows the variation of M_s, M_r, and H_c for the four CFO substituted samples. The increase in M_s, M_r, and H_c with increase in CFO concentration indicate an improvement in short-range ferromagnetic interaction in the ceramics. BNC15 sample shows the highest $2M_s$ (17.84 emu/g), $2M_r$ (8.94 emu/g), and $2H_c$ (2.72 kOe) among the four CFO substituted sample, which is mainly attributed to highly magnetic CFO phase. Moreover, the increase in coercive magnetic field with CFO substitution is ascribed to enhancement in magneto-crystalline anisotropy. The obvious increase is due to increase in number of Co^{2+} ions, which acts as trigger for increasing magneto-crystalline anisotropy. The improvement in magnetic characteristics is also due to increase in domain wall pinning by the highly ferroelectric BCT and NBT phases, which change the magnetization orientation of easy axis (1 0 0) to the hard axis (1 1 1). Further, the Fe ions in CFO enhances the exchange interaction and subsequently supports magnetic linkages, which is responsible for enhancing domain wall motion and hence a higher coercive field is obtained in the ceramics with high CFO concentration [40, 41]. The emergence of magnetic characteristics along with moderately high electrical characteristics in CFO substituted ceramics may result in higher value of ME coefficient (discussed in section 3.8). The field cooled (FC) and zero field cooled (ZFC) magnetization of the four CFO substituted samples were carried out as a function of temperature in the presence of magnetic field. The measurement has been carried out in a wide temperature range of 50 to 400 k in the presence of 12 kOe magnetic field with constant heating rate. It should be noted that the samples with higher CFO concentration have shown significant magnetization behavior even beyond the room temperature as shown in insets of Fig. 4(a-d). Both FC and ZFC loops overlap with each other at higher operating temperature; however, at lower temperature, the FC and ZFC loops separates with each other, which corresponds to their blocking temperature (T_B). It can be observed that all the samples have shown slightly

0018-9464 (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 December 06,2021 at 06:37:35 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/TMAG.2021.3131511, IEEE Transactions on Magnetics

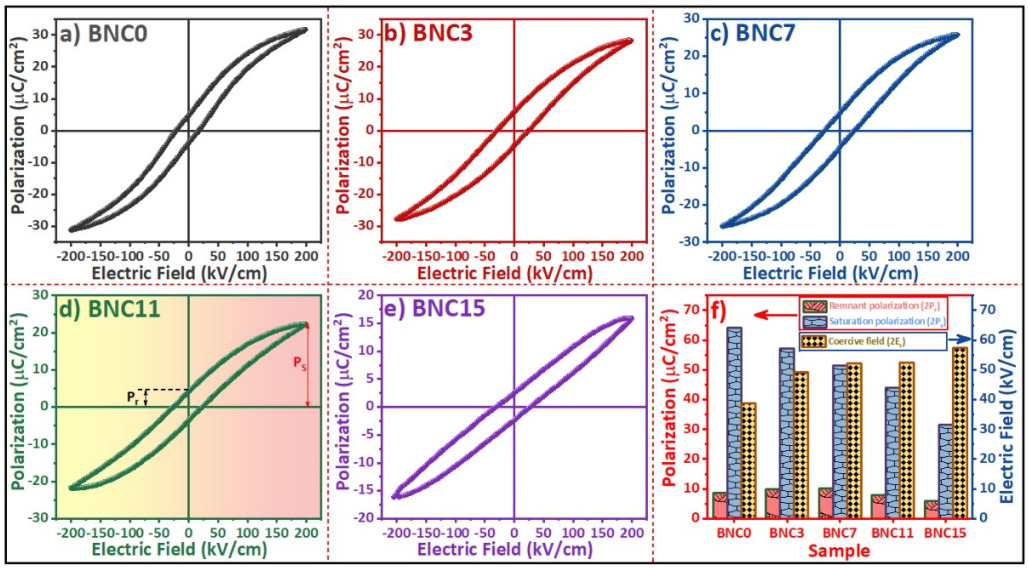


Fig. 6. Room temperature ferroelectric (P-E loops) characteristics for (0.95-*x*)BCT-0.05NBT-*x*CFO ceramic samples: (a) BNC0, (b) BNC3, (c) BNC7, (d) BNC11, and (e) BNC15. (f) Comparison of remnant polarization ($2P_r$), saturation polarization ($2P_S$), and coercive field ($2E_c$) for (0.95-*x*)BCT-0.05NBT-*x*CFO samples.

different T_B. The separation phenomena between FC and ZFC curve originate due to change in thermal energy and a competition between magneto crystalline anisotropy related energy barriers.

D. Ferroelectric investigation

The ferroelectric hysteresis (P-E) loops of (0.95-*x*)BCT-0.05NBT-*x*CFO ceramics are presented in Fig. 6(a-e). It can be observed that the BNC0, BNC3, BNC7, and BNC11 ceramic samples have shown a relaxor ferroelectric character, which is primarily ascribed to the NBT phase. However, at high CFO concentration (BNC15), the loop has started to become ovel shaped, which is not desirable for obtaining superior energy storage performance. Generally, pure $BaTiO_3$ exhibit a classical ferroelectric behavior with high remnant and saturation polarization. In order to improve the energy storage performance, it is required to reduce the remnant polarization as low as possible, which will result in lower loss and subsequently higher discharge energy density and energy efficiency. The substitution of NBT in BCT has resulted in slight reduction in remnant polarization without affecting the saturation polarization of the material, and therefore, in the present work, an optimized composition with BCT and small amount of NBT is considered as the parent composition for preparing a multiferroic ceramic with optimum ferroelectric and ferromagnetic properties. Moreover, it can be emphasized that the addition of NBT in BCT has resulted in classical to ferroelectric transition. The origin of relaxor behavior is mainly ascribed to increase in cation disorder induced by the replacement of Ba^{2+} and Ti^{4+} ions at the A and B-site of $BaTiO_3$ [42, 43]. The difference in the ionic valance and ionic radii results in the weakening of long-range dipole interaction caused by the creation of local polar distortion in the nanoscale regions. A weak coupling between the adjacent clusters, which is induced by the charge difference and composition inhomogeneity is responsible for the formation of these isolated polar nanoregions [44]. Fig. 6(f) depicts the variation of remnant polarization, saturation polarization, and coercive electric field for the five synthesized samples. With increase in CFO concentration, there is a gradual reduction in saturation polarization and found to be minimum for the sample with highest CFO concentration. Further, coercive electric field has shown exactly opposite trend compared to saturation

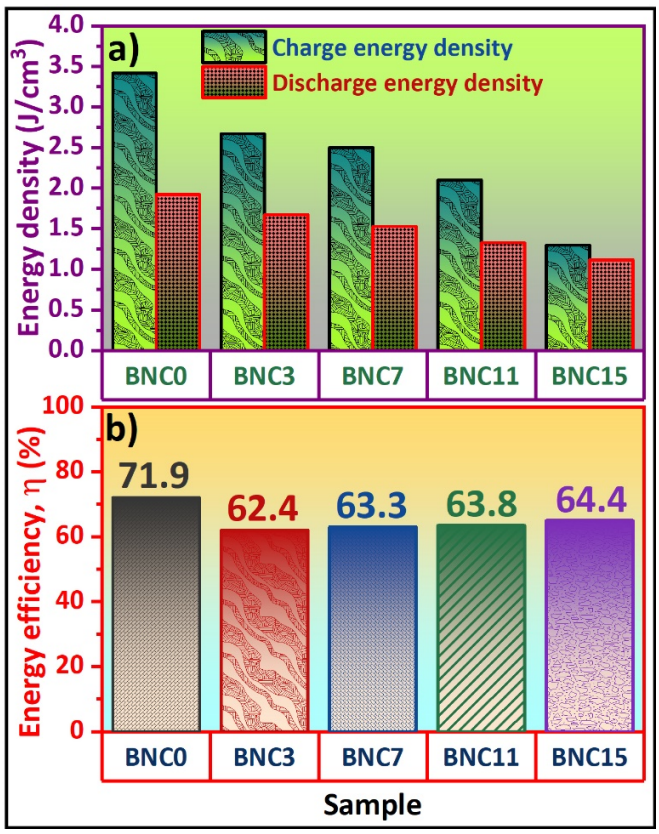


Fig. 7. (a) Charge energy density and discharge (recoverable) energy density for the (0.95-*x*)BCT-0.05NBT-*x*CFO samples. (b) Energy efficiency for the (0.95-*x*)BCT-0.05NBT-*x*CFO samples.

0018-9464 (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 December 06,2021 at 06:37:35 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/TMAG.2021.3131511, IEEE Transactions on Magnetics

Table 2. Comparative analysis of Saturation magnetization, ferroelectric parameters, discharge energy density, energy efficiency, shape factor, and peak magnetoelectric coefficient for the (0.95-x)BCT-0.05NBT-xCFO ceramic samples.

Sample	Saturation Magnetization ($2M_r$) (emu/g)	Saturation polarization ($2P_s$) (μC/cm^2)	ΔP=P_s–P_r (μC/cm^2)	Discharge energy density (at 200 kV/cm) ($W_{discharge}$) (J/cm^3)	Energy efficiency (η)	Shape factor (β)	Peak Magnetoelectric coefficient (α_{ME}) (mV/cm.Oe)
BNC0	-	64.05	27.71	1.924	71.9	13.3	-
BNC3	5.32	57.23	23.58	1.672	62.4	13.4	0.884
BNC7	11.96	51.47	20.62	1.528	63.3	21.5	1.920
BNC11	16.52	44.04	18.08	1.326	63.8	15.9	5.411
BNC15	17.84	31.62	12.88	1.116	64.4	13.1	4.995

polarization, which is primarily ascribed to non-ferroelectric CFO phase. An increase in coercive electric field indicates that a relatively higher electric field is required to polarize the material.

Charge energy density (J_C), discharge (recoverable) energy density (J_D), and energy efficiency (η) has been measure to evaluate the energy storage performance of (0.95-x)BCT-0.05NBT-xCFO ceramics. All these three parameters were calculated by the P-E loop of corresponding ceramic by using the following equations [45]:

$$J_C = \int_0^{P_s} \boldsymbol{E} \cdot \boldsymbol{dP} \quad (1)$$

$$J_D = \int_{P_r}^{P_s} \boldsymbol{E} \cdot \boldsymbol{dP} \quad (2)$$

$$\eta = \frac{J_D}{J_D + Loss} \quad (3)$$

Where, E indicates the applied electric field and P represents the polarization on the application of electric field. The energy storage parameter (J_C, J_D, and η) were evaluated by integrating the net enclosed area between the discharge and charge curve of the hysteresis loops. By observing the Eq. (2), it can be deduced that a larger energy efficiency suggests a smaller energy loss in the form of hysteresis. Fig. 7(a-b) show the variation of energy storage parameters with increase in CFO substitution. It can be observed that the discharge energy density decreases with increase in CFO substitution; however, even at highest CFO concentration, the discharge energy density is high enough. Among the five prepared samples, BNC11 sample has shown highest ME coefficient (discussed in section 3.8), and therefore, the energy storage performance of BNC11 ceramic sample has been evaluated at highest possible applied electric field. A high energy storage performance can be obtained when the difference between saturation polarization and remnant polarization is large. However, the difference ΔP (P_s–P_r) decreases with increase in CFO substitution and consequently, there is slight reduction in energy storage performance. Fig. 8(a) and Fig. 8(b) depict the ferroelectric hysteresis loop and energy storage parameters at a maximum applied electric field of 388 kV/cm. BNC11 sample exhibit a high discharge energy density of 2.249 J/cm^3 and moderately high energy efficiency of 63.8%. It is worth noting that the objective of this investigation is to develop a material with optimum multiferroic characteristics, and therefore, an improvement in magnetic and magnetoelectric characteristics comes at the cost of slight deterioration in electrical properties. However, the energy storage performance obtained in the present work is still high enough, which is attributed to the fact that the electrical properties in (0.95-x)BCT-0.05NBT-xCFO ceramics are still governed by BCT phase.

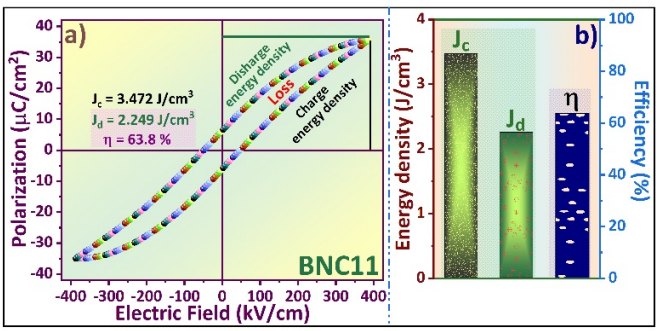


Fig. 8. (a) Ferroelectric hysteresis curve of optimized BNC11 ceramic sample at a maximum applied electric field of 388 kV/cm. (b) Charge energy density (J_c), discharge (recoverable) energy density (J_d), and energy efficiency (η) for the optimized BNC11 sample.

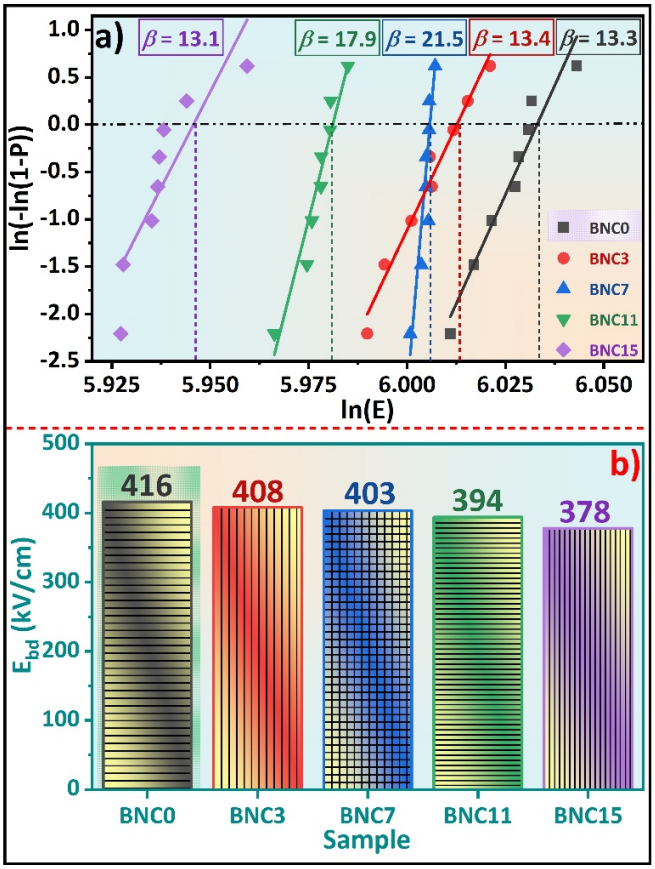


Fig. 9. Comparison of breakdown strength of (0.95-x)BCT-0.05NBT-xCFO ceramic samples evaluated by Weibull distribution function. (b) Experimentally observed breakdown strength for the five synthesized samples.

0018-9464 (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 December 06,2021 at 06:37:35 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/TMAG.2021.3131511, IEEE Transactions on Magnetics

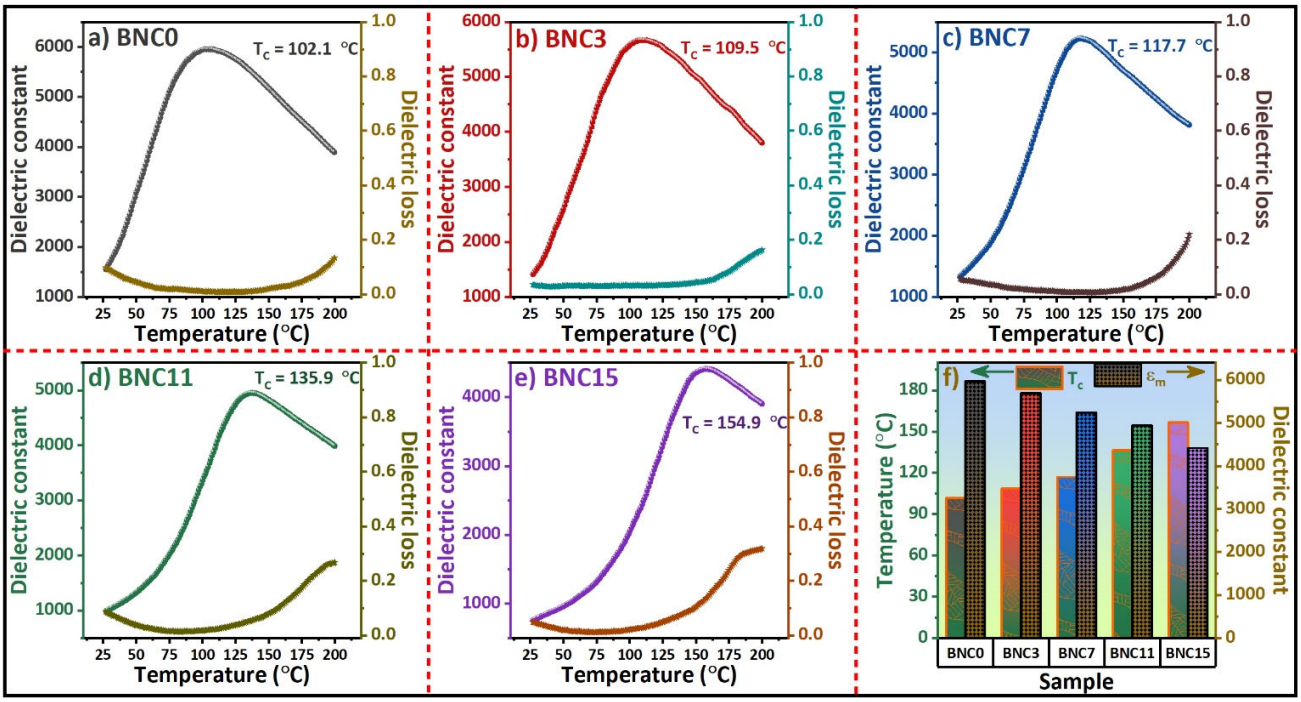


Fig. 10. Temperature dependent dielectric characteristics for (0.95-*x*)BCT-0.05NBT-*x*CFO ceramic samples: (a) BNC0, (b) BNC3, (c) BNC7, (d) BNC11, and (e) BNC15. (f) Comparison of Curie temperature (T_c) and Peak dielectric constant for the (0.95-*x*)BCT-0.05NBT-*x*CFO ceramics.

E. Breakdown strength analysis

Breakdown strength is one of the most important parameters which affect the energy storage performance of ceramics. The breakdown strength of (0.95-*x*)BCT-0.05NBT-*x*CFO ceramics were investigated by using Weibull distribution and shown in Fig. 9(a). The characteristic breakdown strength of prepared ceramics was evaluated by using the following equations [46]:

$$\boldsymbol{X_k = \ln\left(E_k\right)} \tag{4}$$

$$\boldsymbol{Y_k = \ln\left(-\ln\left(1-\frac{k}{1+n}\right)\right)} \tag{5}$$

Where, X_k and Y_k present a linear relationship, k represents the rank of the sample, E_k indicates the breakdown strength of the sample, and n is the number of tested pellet specimens. The Weibull modulus (β) indicates the reliability of experimental data and has been calculated by determining the slope of fitted lines. In general, a β value greater than 1 indicates that the Weibull distribution model is suitable for breakdown strength analysis. The value of β depends on sample uniformity, grain size distribution, and number of defects in the material. In the present work, all the samples have shown a value greater than 13, which indicate the reliability of breakdown strength analysis. Further, BNC7 and BNC11 samples have shown a significantly higher value compared to other three samples, as shown in Fig. 9(b). The substitution of CFO in BCT-NBT has resulted in slight reduction in breakdown strength, which is apparent, as the conductivity of CFO is higher than BCT and NBT. Among the five prepared samples, BNC0 sample exhibit the highest breakdown strength of 416 kV/cm, which could be attributed in small grain size and optimal distribution of BCT and NBT phase. A smaller grain size led to reduction in leakage current and provides a more complex path for the flow of electric current, which allows the material to sustain higher electric field [47]. The introduction of CFO in BCT-NBT ceramic results in slight increase in leakage current density, which subsequently decreases the breakdown strength and the CFO substituted samples breakdown at a relatively lower electric field. The obtained results are in agreement with the previous studies, where also a reduction in breakdown strength has been observed with increase in average grain size [48-50]. However, the CFO substituted sample have still shown relatively high electrical breakdown strength in comparison to various reported magnetoelectric ceramics.

F. Dielectric property analysis

The temperature vs. dielectric constant spectra of (0.95-*x*)BCT-0.05NBT-*x*CFO ceramics is shown in Fig. 10(a-e) at 10 kHz frequency. All the samples have shown a broad phase transition, which may be attributed to breaking of long-range dipolar interaction. With increase in CFO concentration, there is an increase in dielectric relaxation behavior and an apparent diffuse phase transition, which is caused by a difference in valance of cations at A and B site of parent BCT phase, leading to the occurrence of polar nano regions (PNRs) in CFO substituted ceramics [51]. Fig. 10(f) shows the variation of Curie temperature (T_c) and peak dielectric constant (ε_m) with variation in CFO concentration. The substitution of CFO in BCT-NBT ceramic has led to gradual reduction in dielectric constant, which is mainly ascribed to decrease in dipolar interaction [52]. Further, the dielectric anomaly around the phase transition slightly broadens and the Curie temperature (T_c) shifts towards high temperature side, which indicates a slight enhancement in relaxor characteristics. The occurrence of such phenomena is primarily attributed to increase in site

0018-9464 (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 December 06,2021 at 06:37:35 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/TMAG.2021.3131511, IEEE Transactions on Magnetics

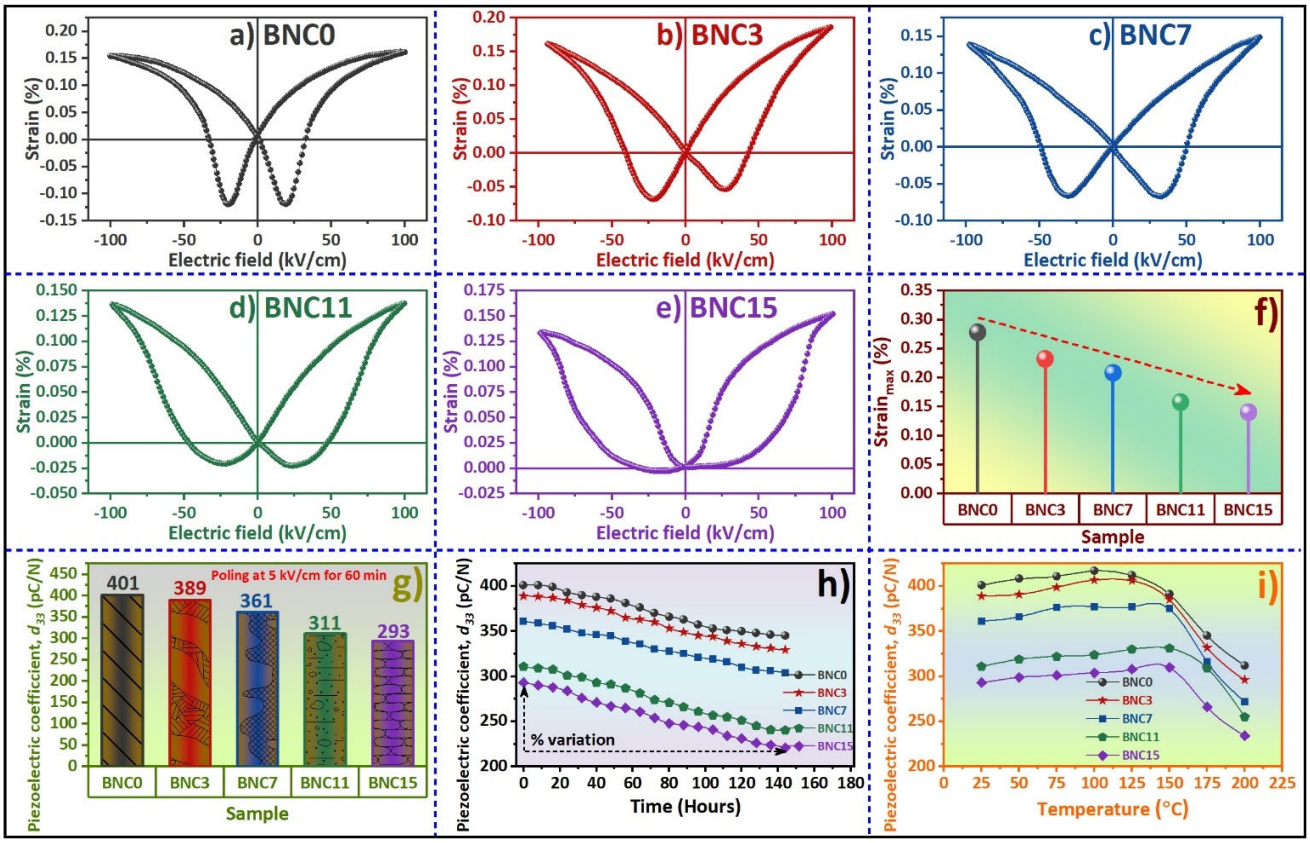


Fig. 11. Room temperature Bipolar strain characteristics for the (0.95-*x*)BCT-0.05NBT-*x*CFO ceramics: (a) BNC0, (b) BNC3, (c) BNC7, (d) BNC11, and (e) BNC15. (f) Variation of percentage maximum strain ($Strain_{max}$) for the five prepared samples. (g) Comparison of room temperature piezoelectric coefficient for the (0.95-*x*)BCT-0.05NBT-*x*CFO ceramic sample. (h) Variation in piezoelectric coefficient after the initial electrical poling. (i) Temperature dependent piezoelectric characteristics for the five synthesized samples.

disorder caused by charge fluctuation, valance state, and vastly different ionic radii of substituent atoms. In order to further explore the relaxor characteristics of (0.95-*x*)BCT-0.05NBT-*x*CFO ceramics, the diffuseness parameter (γ) has been calculated by employing the modified Curie-Weiss law [53]:

$$\frac{1}{\varepsilon_r} - \frac{1}{\varepsilon_m} = \frac{(T-T_m)^\gamma}{C} \quad (6)$$

Where, ε_r represents the relative dielectric constant, ε_m indicates the peak dielectric constant at T_m, and C denotes the Curie-Weiss coefficient. Generally, the γ = 1 indicates a classical ferroelectric character (observed in pure $BaTiO_3$), and a deviation of γ from 1 to 2 suggests a transformation from classical to relaxor ferroelectric. Fig. S4(a-e) in supplementary data depict the diffuseness curve for the corresponding sample. In the present work, BNC0 shows the lowest γ of 1.31 and BNC15 sample shows the highest γ of 1.46, indicating a strong relaxor behavior. These results are ascribed to disordered distribution of Ba^{2+}, Ca^{2+}, Ti^{4+}, Na^{1+}, Bi^{3+}, Co^{2+}, and Fe^{3+} cations in the CFO substituted ceramic, which increases the variation in random local electric field and disrupts the long-range polar ordering of dipoles. The variation trend in dielectric constant with increasing temperature even for high CFO concentration sample indicate that the dielectric characteristics are still governed by parent BCT phase, which has proved to be beneficial for obtaining moderately high dielectric constant along with improved magnetic properties.

The temperature dependent dielectric loss study has also been carried out and shown in Fig. 10(a-e). There is a slight increase in the dielectric loss of the CFO substituted ceramics, which is mainly ascribed to slight increase in leakage current. It can be observed that the dielectric loss is very small in 27- 140 ºC temperature range, which is generally considered as the usable range of various electronic devices. All the samples have shown an abrupt increase in dielectric loss beyond 140-160 ºC (depend on Curie temperature), which is due to increase in thermal activation energy of ions at high operating temperature [54]. Finally, it can be emphasized that the optimized sample (i.e., BNC11) still exhibit moderately high dielectric constant along with low dielectric loss, which will be useful in obtaining improved magnetoelectric characteristics.

G. Electromechanical analysis

To better understand the multiferroic characteristics of (0.95-*x*)BCT-0.05NBT-*x*CFO ceramics, room temperature bipolar strain characteristics and piezoelectric coefficient has been thoroughly analyzed. Fig. 11(a-e) show the S-E curve for the five synthesized samples at a maximum applied electric field of 100 kV/cm and 1 Hz frequency. All the samples exhibit a nearly symmetrical butterfly shaped profile with significant negative strain except for BNC15 sample. Even a small amount of CFO substitution results in considerable change in electric field induced strain characteristics. Further, it can be observed that except for BNC15 sample, all the other four samples exhibit a

0018-9464 (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 December 06,2021 at 06:37:35 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/TMAG.2021.3131511, IEEE Transactions on Magnetics

typical butterfly shaped nonergodic relaxor characteristic. Similar to ferroelectric measurement, with increase in CFO concentration, the loop has become more and more broad and there is gradual reduction in $strain_{max}$. Fig. 11(f) show the variation in $strain_{max}$ with change in CFO concentration. These phenomena indicate that the introduction of CFO in BCT-NBT has led to decrease in overall resistivity of the material. The negative strain tends to decrease with increase in CFO substitution, which may be attributed to disturbance in long-range ferroelectric ordering caused by the enhancement in random fields and structural disorder [55]. Among the five prepared samples, BNC0 exhibit the highest $strain_{max}$ (0.2783%), which is apparent due to inherit strain characteristics of BCT and NBT phase.

Next, the piezoelectric coefficient (D_{33}) of all the samples has been measured and shown in Fig. 11(g). It is worth mentioning that an optimum poling electric field, poling duration, and poling temperature plays a vital role in obtaining superior piezoelectric characteristics [56, 57]. In the present work, to obtain an accurate comparison between the five synthesized samples, the electric poling is carried out at 5 kV/cm for 60 minutes at 100 ºC. It is apparent from P-E and S-E characteristics that the substitution of CFO will result in slight reduction in piezoelectric coefficient. Among the five prepared samples, BNC0 sample shows a piezoelectric coefficient of 401 pC/N and the BNC11 sample with optimum magnetoelectric characteristics exhibit a moderately high piezoelectric coefficient of 311 pC/N. The aging properties of all the samples has been measured and shown in Fig. 11(h). The piezoelectric coefficient of all the samples has been measured after every 8 hours for up to 144 hours. The percentage variation in piezoelectric coefficient were found be slightly higher in BNC11 and BNC15 sample. However, the obtained piezoelectric characteristics are still superior as we have considered a long time frame of 144 hours. Finally, the temperature dependent piezoelectric properties were measured and shown in Fig. 11(i). Interestingly, all the samples have shown a stable thermal characteristic in a wide temperature window (27-150 ºC), which is very beneficial for various microelectronics-based devices. There is approximately 2-3% variation in 27-150 ºC temperature and therefore, the prepared materials can be easily utilized in a wide temperature range for various electromechanical applications. It should also be emphasized that the obtained electromechanical characteristics in the present work are still high enough from the perspective of multiferroic material.

H. Magnetoelectric study

The magnetoelectric study of CFO substituted samples have been carried out and shown Fig. 12(a-c). In the present study, the magnetoelectric characteristics of BNC0 sample has not been shown due to insignificant magnitude. In the present investigation, for the measurement of ME coefficient, an in-house dynamic technique has been used, in which an AC magnetic field is superimposed on to a DC magnetic field. For the measurement, the samples were electrically poled at 5 kV/cm for 60 minutes in order to get finite electric polarization. Further, during the ME measurement, the poled pellet samples were kept perpendicular to the magnetic field direction. The application of magnetic field results in the induction of voltage due to magnetoelectric effect, which is subsequently measured by a lock-in amplifier. Generally, strain mediated coupling in multiferroic ceramics and intrinsic effects in single phase ceramics are responsible of the emergence of magnetoelectric coupling [58-60]. In case of ceramics, on the application of magnetic field, the magnetic phase changes its dimension due to magnetostriction effect [61, 62]. The strain generated in the magnetic phase (CFO) is subsequently transferred to the dielectric/ferroelectric phase (BCT), which results in changes in ferroelectric characteristics. The magnetoelectric study has been carried out to determine the coupling between ferroelectric BCT-NBT and magnetic CFO phase. First, the frequency dependent magnetoelectric coefficient (α_{ME}) has been measured at AC magnetic field (H_{AC}) of 15 Oe and DC magnetic field (H_{DC}) of 4 kOe. The α_{ME} increases with increasing CFO concentration up to x = 0.11 and then decreased on further increasing the CFO concentration. The BNC11 sample has shown a maximum α_{ME} of 5.411 mV/cm.Oe at 12.15 kHz frequency. Next, the variation of α_{ME} has been measured as a function of H_{DC} at the optimized frequency of 12 kHz and H_{AC} of 15 Oe. Again, the BNC11 has shown the highest α_{ME} among the four CFO substituted samples and the maximum was found at H_{DC} of ~4.21 kOe. The coexistence of strain mediated electric-mechanical-magnetic energy conversion at the ferroelectric, piezoelectric, and magnetostrictive interfaces has given rise to substantial ME effect [63]. The results obtained in the present work are in complete agreement with the previous reports on $BaTiO_3$-$CoFe_2O_4$ based ceramics. Wadgane et al.

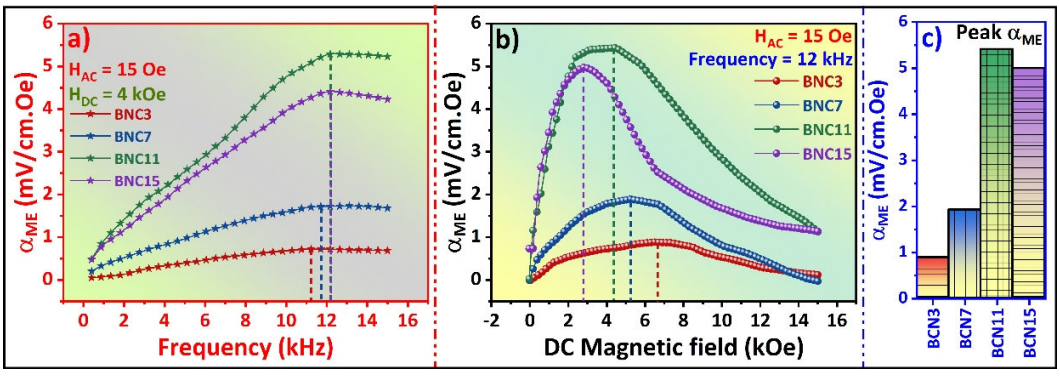


Fig. 12. (a) Variation of magnetoelectric coefficient (α_{ME}) with frequency for the CFO substituted samples. (b) Variation of α_{ME} with change in DC magnetic field (observed at the optimized frequency of 12 kHz). (c) Comparison of peak α_{ME} for the four CFO substituted samples.

0018-9464 (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 December 06,2021 at 06:37:35 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/TMAG.2021.3131511, IEEE Transactions on Magnetics

investigated the magnetoelectric characteristics of $Na_{0.5}Bi_{0.5}TiO_3$–$Co_{0.75}Zn_{0.25}Cr_{0.2}Fe_{1.8}O_4$ ceramics and shown an improvement in magnetoelectric properties [64]. In another report, Gaikwad et al. synthesized $Co_{0.7}Fe_{2.3}O_4$-$SrTiO_3$ nano-composites and observed a significant improvement in magnetoelectric characteristics [17]. Further, it has been observed that there is a limit of up to which magnetoelectric coefficient increases with increasing magnetic field and after that limit, it starts to decrease [18]. It should be noted that α_{ME} value has been substantially improved, which may be attributed to enhancement in bidirectional magnetostriction. In addition to this, the value of α_{ME} of ceramics also depends on synthesis method, final sintering temperature, grain size, grain agglomeration, and type of interface between the electrical and magnetic phases. Finally, it can be emphasized that the suitable amount of CFO substitution in ferroelectric BCT-NBT ceramic has a significant effect in improving ME characteristics and therefore, the optimized BNC11 ceramic with excellent magnetic, electrical, and magnetoelectric characteristics can be utilized in various multifunctional applications.

IV. Conclusions

Novel $(0.95\text{-}x)Ba_{0.9}Ca_{0.1}TiO_3$-$0.05Na_{0.5}Bi_{0.5}TiO_3$-$xCoFe_2O_4$ ceramics were successfully synthesized via mechanochemical activation method. Rietveld investigation confirmed the presence of tetragonal (BCT), rhombohedral (NBT), and cubic phase (CFO) for the CFO substituted samples. The substitution of CFO in BCT-NBT has led to improvement in magnetic and magnetoelectric properties. There is a slight decrease in ferroelectric, dielectric, and electromechanical characteristics of CFO substituted samples in comparison to parent phase (BCT-NBT), which is primarily attributed to relatively higher conductivity of CFO compared to BCT and NBT. However, the electrical properties of CFO substituted samples are still high enough, which has led to substantial improvement in magnetoelectric characteristics of the material. The sample with x = 0.11 CFO concentration has shown the highest magnetoelectric coefficient (5.411 mV/cm.Oe) among the five synthesized samples. Further, the sample with x = 0.11 CFO concentration large discharge energy density (2.249 J/cm^3) and moderately high energy efficiency (63.8%) at the maximum applied electric field of 388 kV/cm. The emergence of magnetic properties along with moderately high electrical characteristics in CFO substituted ceramics suggests their applicability in the design and development of multifunctional devices.

References

[1] N. Ortega, A. Kumar, J. F. Scott, and R. S. Katiyar, "Multifunctional magnetoelectric materials for device applications," *Journal of Physics: Condensed Matter,* vol. 27, p. 504002, 2015/11/27 2015.

[2] M. PourhosseiniAsl, X. Gao, S. Kamalisiahroudi, Z. Yu, Z. Chu, J. Yang, *et al.*, "Versatile power and energy conversion of magnetoelectric composite materials with high efficiency via electromechanical resonance," *Nano Energy,* vol. 70, p. 104506, 2020/04/01/ 2020.

[3] N. Kumar, A. Shukla, N. Kumar, and R. N. P. Choudhary, "Structural, electrical and magnetic properties of eco-friendly complex multiferroic material: $Bi(Co_{0.35}Ti_{0.35}Fe_{0.30})O_3$," *Ceramics International,* vol. 45, pp. 822-831, 2019/01/01/ 2019.

[4] D. Lisjak and A. Mertelj, "Anisotropic magnetic nanoparticles: A review of their properties, syntheses and potential applications," *Progress in Materials Science,* vol. 95, pp. 286-328, 2018/06/01/ 2018.

[5] A. Jain, Y. Wang, N. Wang, Y. Li, and F. Wang, "Tuning the dielectric, ferroelectric and electromechanical properties of $Ba_{0.83}Ca_{0.10}Sr_{0.07}TiO_3$–$MnFe_2O_4$ multiferroic composites," *Ceramics International,* vol. 46, pp. 7576-7585, 2020.

[6] M. Arshad, H. Du, M. S. Javed, A. Maqsood, I. Ashraf, S. Hussain, *et al.*, "Fabrication, structure, and frequency-dependent electrical and dielectric properties of Sr-doped $BaTiO_3$ ceramics," *Ceramics International,* vol. 46, pp. 2238-2246, 2020/02/01/ 2020.

[7] K. Tsuji, A. Ndayishimiye, S. Lowum, R. Floyd, K. Wang, M. Wetherington, *et al.*, "Single step densification of high permittivity $BaTiO_3$ ceramics at 300 ºC," *Journal of the European Ceramic Society,* vol. 40, pp. 1280-1284, 2020/04/01/ 2020.

[8] A. Wohninsland, A.-K. Fetzer, A. Riaz, H.-J. Kleebe, J. Rödel, and L. K. Venkataraman, "Correlation between enhanced lattice distortion and volume fraction of polar nanoregions in quenched $Na_{1/2}Bi_{1/2}TiO_3$–$BaTiO_3$ ceramics," *Applied Physics Letters,* vol. 118, p. 072903, 2021.

[9] E. Haily, L. Bih, A. El bouari, A. Lahmar, M. Elmarssi, and B. Manoun, "Effect of BaO–Bi_2O_3–P_2O_5 glass additive on structural, dielectric and energy storage properties of $BaTiO_3$ ceramics," *Materials Chemistry and Physics,* vol. 241, p. 122434, 2020/02/01/ 2020.

[10] A. Jain, Y. G. Wang, N. Wang, and F. L. Wang, "Critical role of CuO doping on energy storage performance and electromechanical properties of $Ba_{0.8}Sr_{0.1}Ca_{0.1}Ti_{0.9}Zr_{0.1}O_3$ ceramics," *Ceramics International,* vol. 46, pp. 18800-18812, 2020/08/01/ 2020.

[11] Y. Yang, H. Hao, L. Zhang, C. Chen, Z. Luo, Z. Liu, *et al.*, "Structure, electrical and dielectric properties of Ca substituted $BaTiO_3$ ceramics," *Ceramics International,* vol. 44, pp. 11109-11115, 2018/07/01/ 2018.

[12] X. Cheng and M. Shen, "Enhanced spontaneous polarization in Sr and Ca co-doped $BaTiO_3$ ceramics," *Solid State Communications,* vol. 141, pp. 587-590, 2007/03/01/ 2007.

[13] Y. Sun, H. Liu, H. Hao, L. Zhang, and S. Zhang, "The role of Co in the $BaTiO_3$–$Na_{0.5}Bi_{0.5}TiO_3$ based X9R ceramics," *Ceramics International,* vol. 41, pp. 931-939, 2015/01/01/ 2015.

[14] C. Yang, P. Lv, J. Qian, Y. Han, J. Ouyang, X. Lin, *et al.*, "Fatigue-Free and Bending-Endurable Flexible Mn-Doped $Na_{0.5}Bi_{0.5}TiO_3$-$BaTiO_3$-$BiFeO_3$ Film Capacitor with an Ultrahigh Energy Storage Performance," *Advanced Energy Materials,* vol. 9, p. 1803949, 2019.

[15] W. Ge, C. Luo, Q. Zhang, Y. Ren, J. Li, H. Luo, *et al.*, "Evolution of structure in $Na_{0.5}Bi_{0.5}TiO_3$ single crystals with $BaTiO_3$," *Applied Physics Letters,* vol. 105, p. 162913, 2014.

[16] A. Jain, Y. Wang, H. Guo, and N. Wang, "Grain size engineered $Ba_{0.9}Sr_{0.1}Ti_{0.9}Hf_{0.1}O_3$-$Na_{0.5}Bi_{0.5}TiO_3$ relaxor ceramics with improved energy storage performance," *Journal of the American Ceramic Society,* vol. 103, pp. 6308-6318, 2020.

[17] A. S. Gaikwad, R. H. Kadam, S. E. Shirsath, S. R. Wadgane, J. Shah, R. K. Kotnala, *et al.*, "Surprisingly high magneto-electric coupling in cubic $Co_{0.7}Fe_{2.3}O_4$-$SrTiO_3$ nano-composites," *Journal of Alloys and Compounds,* vol. 773, pp. 564-570, 2019/01/30/ 2019.

[18] A. S. Gaikwad, S. E. Shirsath, S. R. Wadgane, R. H. Kadam, J. Shah, R. K. Kotnala, *et al.*, "Magneto-electric coupling and improved dielectric constant of $BaTiO_3$ and Fe-rich ($Co_{0.7}Fe_{2.3}O_4$) ferrite nano-composites," *Journal of Magnetism and Magnetic Materials,* vol. 465, pp. 508-514, 2018/11/01/ 2018.

[19] S. E. Shirsath, X. Liu, M. Assadi, A. Younis, Y. Yasukawa, S. K. Karan, *et al.*, "Au quantum dots engineered room temperature crystallization and magnetic anisotropy in $CoFe_2O_4$ thin films," *Nanoscale horizons,* vol. 4, pp. 434-444, 2019.

[20] S. E. Shirsath, D. Wang, J. Zhang, A. Morisako, S. Li, and X. Liu, "Single-Crystal-like Textured Growth of $CoFe_2O_4$ Thin Film on an Amorphous Substrate: A Self-Bilayer Approach," *ACS Applied Electronic Materials,* vol. 2, pp. 3650-3657, 2020/11/24 2020.

[21] S. E. Shirsath, X. Liu, Y. Yasukawa, S. Li, and A. Morisako, "Switching of magnetic easy-axis using crystal orientation for large perpendicular coercivity in $CoFe_2O_4$ thin film," *Scientific Reports,* vol. 6, p. 30074, 2016/07/20 2016.

[22] Y. Zhang, T. Lu, Y. Ye, W. Dai, Y. a. Zhu, and Y. Pan, "Stabilizing Oxygen Vacancy in Entropy-Engineered $CoFe_2O_4$-Type Catalysts for Co-prosperity of Efficiency and Stability in an Oxygen Evolution Reaction," *ACS Applied Materials & Interfaces,* vol. 12, pp. 32548-32555, 2020/07/22 2020.

[23] H. Zheng, J. Wang, S. E. Lofland, Z. Ma, L. Mohaddes-Ardabili, T. Zhao, *et al.*, "Multiferroic $BaTiO_3$-$CoFe_2O_4$ Nanostructures," *Science,* vol. 303, p. 661, 2004.

[24] A. Jain, Y. G. Wang, N. Wang, Y. Li, and F. L. Wang, "Existence of heterogeneous phases with significant improvement in electrical and

0018-9464 (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 December 06,2021 at 06:37:35 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/TMAG.2021.3131511, IEEE Transactions on Magnetics

magnetoelectric properties of $BaFe_{12}O_{19}/BiFeO_3$ multiferroic ceramic composites," *Ceramics International,* vol. 45, pp. 22889-22898, 2019/12/01/ 2019.

[25] S. S. Choudhari, S. B. Shelke, K. M. Batoo, S. F. Adil, A. B. Kadam, A. Imran*, et al.*, "$Mn_{0.7}Zn_{0.3}Fe_2O_4 + BaTiO_3$ composites: structural, morphological, magnetic, M–E effect and dielectric properties," *Journal of Materials Science: Materials in Electronics,* vol. 32, pp. 10308-10319, 2021/04/01 2021.

[26] S. S. Choudhari, S. R. Wadgane, B. P. Gaikwad, S. S. Satpute, K. M. Batoo, O. M. Aldossary*, et al.*, "Strain mediated enhancement in magnetoelectric properties of sonochemically synthesized piezoelectric and piezomagnetic composites," *Ceramics International,* vol. 47, pp. 6496-6504, 2021/03/01/ 2021.

[27] S. R. Wadgane, S. E. Shirsath, A. S. Gaikwad, S. Satpute, A. B. Kadam, and R. H. Kadam, "Ferro- and magneto-electric characteristics in $YFeO_3$−$Y_3Fe_5O_{12}$ nanocomposites," *Journal of Magnetism and Magnetic Materials,* vol. 457, pp. 103-109, 2018/07/01/ 2018.

[28] R. Gao, Z. Wang, G. Chen, X. Deng, W. Cai, and C. Fu, "Influence of core size on the multiferroic properties of $CoFe_2O_4$@$BaTiO_3$ core shell structured composites," *Ceramics International,* vol. 44, pp. S84-S87, 2018/11/01/ 2018.

[29] N. Jedrecy, T. Aghavnian, J.-B. Moussy, H. Magnan, D. Stanescu, X. Portier*, et al.*, "Cross-Correlation between Strain, Ferroelectricity, and Ferromagnetism in Epitaxial Multiferroic $CoFe_2O_4/BaTiO_3$ Heterostructures," *ACS Applied Materials & Interfaces,* vol. 10, pp. 28003-28014, 2018/08/22 2018.

[30] Z. Wang, R. Gao, G. Chen, X. Deng, W. Cai, and C. Fu, "Dielectric, ferroelectric and magnetoelectric properties of in-situ synthesized $CoFe_2O_4/BaTiO_3$ composite ceramics," *Ceramics International,* vol. 46, pp. 9154-9160, 2020/05/01/ 2020.

[31] S. Ahmed, M. Atif, A. U. Rehman, S. Bashir, N. Iqbal, W. Khalid*, et al.*, "Enhancement in the magnetoelectric and energy storage properties of core-shell-like $CoFe_2O_4$−$BaTiO_3$ multiferroic nanocomposite," *Journal of Alloys and Compounds,* p. 160875, 2021/06/23/ 2021.

[32] K. Shahzad, S. Mushtaq, M. Rizwan, W. Khalid, M. Atif, F. U. Din*, et al.*, "Field-controlled magnetoelectric core-shell $CoFe_2O_4$@$BaTiO_3$ nanoparticles as effective drug carriers and drug release in vitro," *Materials Science and Engineering: C,* vol. 119, p. 111444, 2021/02/01/ 2021.

[33] A. Jain, A. K. Panwar, and A. K. Jha, "Significant improvement in morphological, dielectric, ferroelectric and piezoelectric characteristics of $Ba_{0.9}Sr_{0.1}Ti_{0.9}Zr_{0.1}O_3$–$BaNb_2O_6$ nanocomposites," *Journal of Materials Science: Materials in Electronics,* vol. 29, pp. 19086-19098, 2018/11/01 2018.

[34] A. Jain, A. K. Panwar, and A. K. Jha, "Structural, dielectric and ferroelectric studies of $Ba_{1-x}Sr_xTiO_3$ ceramics prepared by mechanochemical activation technique," *Journal of Materials Science: Materials in Electronics,* vol. 27, pp. 9911-9919, 2016/09/01 2016.

[35] V. Buscaglia and C. A. Randall, "Size and scaling effects in barium titanate. An overview," *Journal of the European Ceramic Society,* vol. 40, pp. 3744-3758, 2020/09/01/ 2020.

[36] J. M. D. Coey, A. P. Douvalis, C. B. Fitzgerald, and M. Venkatesan, "Ferromagnetism in Fe-doped SnO_2 thin films," *Applied Physics Letters,* vol. 84, pp. 1332-1334, 2004.

[37] A. Jain and A. K. Panwar, "Synergetic effect of rare-earths doping on the microstructural and electrical properties of Sr and Ca co-doped $BaTiO_3$ nanoparticles," *Ceramics International,* vol. 46, pp. 10270-10278, 2020/06/01/ 2020.

[38] J. Lee, W. D. Lu, and E. Kioupakis, "Electronic and optical properties of oxygen vacancies in amorphous Ta_2O_5 from first principles," *Nanoscale,* vol. 9, pp. 1120-1127, 2017.

[39] H. Chen, C. S. Tu, T. Wang, Y. Ding, T. C. Liao, L. S. Jou*, et al.*, "Synthesis and Magnetic Properties of $BiFeO_3$, $(BiFeO_3)_{0.95}(BaTiO_3)_{0.05}$, and $(Bi_{0.95}Ba_{0.05})(Fe_{0.95}Ti_{0.05})O_3$ Ceramics," *IEEE Transactions on Magnetics,* vol. 47, pp. 706-709, 2011.

[40] X. Wang, Y. Lu, T. Zhu, S. Chang, and W. Wang, "$CoFe_2O_4$/N-doped reduced graphene oxide aerogels for high-performance microwave absorption," *Chemical Engineering Journal,* vol. 388, p. 124317, 2020/05/15/ 2020.

[41] A. El Arrassi, Z. Liu, M. V. Evers, N. Blanc, G. Bendt, S. Saddeler*, et al.*, "Intrinsic Activity of Oxygen Evolution Catalysts Probed at Single $CoFe_2O_4$ Nanoparticles," *Journal of the American Chemical Society,* vol. 141, pp. 9197-9201, 2019/06/12 2019.

[42] H. Zhao, X. Yang, D. Pang, and X. Long, "Enhanced energy storage efficiency by modulating field-induced strain in $BaTiO_3$-$Bi(Ni_{2/3}Ta_{1/3})O_3$ lead-free ceramics," *Ceramics International,* 2021/05/14/ 2021.

[43] A. Jain, Y. G. Wang, N. Wang, Y. Li, and F. L. Wang, "Synergetic effect of $BiYb_{0.9}Sc_{0.1}O_3$ substitution on the energy storage performance of $Ba_{0.90}Sr_{0.10}Ti_{0.90}Zr_{0.10}O_3$ ferroelectric ceramic," *Journal of Physics D: Applied Physics,* vol. 53, p. 235301, 2020/04/03 2020.

[44] V. V. Shvartsman and D. C. Lupascu, "Lead-Free Relaxor Ferroelectrics," *Journal of the American Ceramic Society,* vol. 95, pp. 1-26, 2012.

[45] W. Huang, Y. Chen, X. Li, G. Wang, N. Liu, S. Li*, et al.*, "Ultrahigh recoverable energy storage density and efficiency in barium strontium titanate-based lead-free relaxor ferroelectric ceramics," *Applied Physics Letters,* vol. 113, p. 203902, 2018.

[46] A. E. L. B. A. Ahmad and M. G. M. Ghazal, "Exponentiated additive Weibull distribution," *Reliability Engineering & System Safety,* vol. 193, p. 106663, 2020/01/01/ 2020.

[47] M. Li, P. Fan, W. Ma, K. Liu, J. Zang, C. Samart*, et al.*, "Constructing layered structures to enhance the breakdown strength and energy density of $Na_{0.5}Bi_{0.5}TiO_3$-based lead-free dielectric ceramics," *Journal of Materials Chemistry C,* vol. 7, pp. 15292-15300, 2019.

[48] S. Li, H. Nie, G. Wang, C. Xu, N. Liu, M. Zhou*, et al.*, "Significantly enhanced energy storage performance of rare-earth-modified silver niobate lead-free antiferroelectric ceramics via local chemical pressure tailoring," *Journal of Materials Chemistry C,* vol. 7, pp. 1551-1560, 2019.

[49] Z. Yang, H. Du, L. Jin, Q. Hu, H. Wang, Y. Li*, et al.*, "Realizing high comprehensive energy storage performance in lead-free bulk ceramics via designing an unmatched temperature range," *Journal of Materials Chemistry A,* vol. 7, pp. 27256-27266, 2019.

[50] A. Jain, Y. G. Wang, N. Wang, Y. Li, and F. L. Wang, "Emergence of ferrimagnetism along with magnetoelectric coupling in $Ba_{0.83}Sr_{0.07}Ca_{0.10}TiO_3/BaFe_{12}O_{19}$ multiferroic composites," *Journal of Alloys and Compounds,* vol. 818, p. 152838, 2020/03/25/ 2020.

[51] S. Praharaj, D. Rout, S. Anwar, and V. Subramanian, "Polar nano regions in lead free $(Na_{0.5}Bi_{0.5})TiO_3$-$SrTiO_3$-$BaTiO_3$ relaxors: An impedance spectroscopic study," *Journal of Alloys and Compounds,* vol. 706, pp. 502-510, 2017/06/05/ 2017.

[52] X. Xie, C. Yang, X.-d. Qi, J.-h. Yang, Z.-w. Zhou, and Y. Wang, "Constructing polymeric interlayer with dual effects toward high dielectric constant and low dielectric loss," *Chemical Engineering Journal,* vol. 366, pp. 378-389, 2019/06/15/ 2019.

[53] W. Kirsch and G. Toth, "Two Groups in a Curie–Weiss Model with Heterogeneous Coupling," *Journal of Theoretical Probability,* vol. 33, pp. 2001-2026, 2020/12/01 2020.

[54] S. Ma, M. Jiang, P. Tao, C. Song, J. Wu, J. Wang*, et al.*, "Temperature effect and thermal impact in lithium-ion batteries: A review," *Progress in Natural Science: Materials International,* vol. 28, pp. 653-666, 2018/12/01/ 2018.

[55] N. S. Sowmya, P. Varade, N. Venkataramani, and A. R. Kulkarni, "Lead-free $Na_{0.4}K_{0.1}Bi_{0.5}TiO_3$ ceramic: Poling effect and enhancement in electromechanical and piezoelectric voltage coefficient," *Journal of the European Ceramic Society,* vol. 40, pp. 5384-5391, 2020/12/01/ 2020.

[56] H. Mana-ay, C.-Y. Wang, K.-Y. Hung, P.-Y. Chen, C.-S. Tu, and C.-S. Chen, "Electric field poling effect on the photosensitivity of samarium-doped bismuth ferrite ceramics," *Ceramics International,* vol. 47, pp. 12574-12582, 2021/05/01/ 2021.

[57] A. Jain, A. K. Panwar, and A. K. Jha, "Influence of milling duration on microstructural, electrical, ferroelectric and piezoelectric properties of $Ba_{0.9}Sr_{0.1}Zr_{0.04}Ti_{0.96}O_3$ ceramic," *Ceramics International,* vol. 42, pp. 18771-18778, 2016/12/01/ 2016.

[58] K. Bhoi, H. S. Mohanty, Ravikant, M. F. Abdullah, D. K. Pradhan, S. N. Babu*, et al.*, "Unravelling the nature of magneto-electric coupling in room temperature multiferroic particulate $(PbFe_{0.5}Nb_{0.5}O_3)$–$(Co_{0.6}Zn_{0.4}Fe_{1.7}Mn_{0.3}O_4)$ composites," *Scientific Reports,* vol. 11, p. 3149, 2021/02/04 2021.

[59] G. F. Dionne, C. Guang-Ji, D. T. McAvoy, B. L. Halpern, and J. J. Schmitt, "Magnetic and stress characterization of nickel ferrite ceramic films grown by jet vapor deposition," *IEEE Transactions on Magnetics,* vol. 31, pp. 3853-3855, 1995.

[60] N. Wang, A. Jain, Y. Li, F. L. Wang, Y. L. Lu, H. Zhen*, et al.*, "Investigation of structural, ferroelectric and magnetic properties of Ca modified $BiFeO_3$–$BaTiO_3$ ceramics," *Ceramics International,* vol. 46, pp. 3855-3860, 2020/02/15/ 2020.

[61] P. N. Anantharamaiah, H. M. Shashanka, R. Kumar, J. A. Chelvane, and B. Sahoo, "Chemically enabling $CoFe_2O_4$ for magnetostrictive strain sensing applications at lower magnetic fields: Effect of Zn substitution,"

0018-9464 (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 December 06,2021 at 06:37:35 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/TMAG.2021.3131511, IEEE Transactions on Magnetics

Materials Science and Engineering: B, vol. 266, p. 115080, 2021/04/01/ 2021.

[62] A. Jain, N. Maikhuri, R. Saroha, M. Pastor, A. Jha, and A. Panwar, "Microstructural and dielectric investigations of vanadium substituted barium titanate ceramics," *Advanced Material Letters,* vol. 7, pp. 567-572, 2016.

[63] Y. Lu, R. Fei, X. Lu, L. Zhu, L. Wang, and L. Yang, "Artificial Multiferroics and Enhanced Magnetoelectric Effect in van der Waals Heterostructures," *ACS Applied Materials & Interfaces,* vol. 12, pp. 6243-6249, 2020/02/05 2020.

[64] S. R. Wadgane, S. T. Alone, A. Karim, G. Vats, S. E. Shirsath, and R. H. Kadam, "Magnetic field induced polarization and magnetoelectric effect in $Na_{0.5}Bi_{0.5}TiO_3$–$Co_{0.75}Zn_{0.25}Cr_{0.2}Fe_{1.8}O_4$ multiferroic composite," *Journal of Magnetism and Magnetic Materials,* vol. 471, pp. 388-393, 2019/02/01/ 2019.

Aditya Jain (M'19) is currently an Assistant Professor with Electronics and Telecommunication Engineering Department, Symbiosis Institute of Technology, Pune, India. He has authored or co-authored 37 research articles in various reputed international journals and conferences. Dr. Jain is a member of IEEE and American Ceramic Society.

Ajay Kumar (SM'19) is currently an Assistant Professor with the Department of Electronics and Communication, Jaypee Institute of Information Technology, Noida, India. He has authored or co-authored more than 60 papers in various reputed international journals and conferences. Dr. Kumar is a Senior Member of IEEE and Member of OSA.

Neha Gupta (M'19) received her B.Sc. (H) and M.Sc. degrees in Electronics from the University of Delhi, India, in 2010 and 2012 respectively. She received her Ph.D. degree from Delhi Technological University in Applied Physics in 2017. She is currently working as an Assistant Professor in Department of Applied Science and Humanities (Physics), Northern India Engineering College, New Delhi, India. She has authored or co-authored more than 25 papers in various reputed international and national journals and conferences.

0018-9464 (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 December 06,2021 at 06:37:35 UTC from IEEE Xplore. Restrictions apply.

https://doi.org/10.36375/prepare_u.iei.a221

SSR Stability Analysis and Modeling for DFIG Connected to Series Compensated Transmission Line

Neevatika Verma, Narendra Kumar
Department of Electrical Engineering
(Delhi Technological University)
(New Delhi-110042)
{nivetikaverma2311@gmail.com,}

Abstract - Due to the massive increase in electrical energy demand, wind power penetration into electric networks is increasing. Usually, the wind farms are located away from the grid, necessitating a series compensated transmission system. Series compensation is always prone to sub-synchronous resonance (SSR), which may lead to the torsional interactions resulting T-G shaft fatigue and associated protection system. In this paper, a modified IEEE first Benchmark Model is used for SSR analysis. A doubly-fed induction generator (DFIG) has been considered supplying power to the grid through a series compensated transmission line. The dynamics of the Rotor Side Converter (RSC) and the Grid Side Converter (GSC) is used to model the DFIG. Thus, it can generate a significant impact in system dynamics. However, previous works have considered a constant voltage source model for RSC and GSC. Therefore, consequently these models are not capable of replicating the dynamic impact of SSR for some critical factors. The inclusion of GSC and RSC modeling in analysis reflects dynamics that are more accurate. The eigenvalue analysis has been used to study the impact of SSR under different operating situations, for instance, wind speed fluctuations and varying the series compensation. To validate the analysis, time-domain simulations have been carried out by using MATLAB.

Keywords – Doubly Fed induction Generator (DFIG), Series Compensation, Sub-synchronous Resonance (SSR), Stability

INTRODUCTION

As renewable energy is the current requirement to meet the global power need without much impact on the environment. Wind power plays a significant role in fulfilling the power requirement. To accommodate the additional power, transmission infrastructural changes are frequently required for wind farms located far from load centers. From the economic considerations, series capacitor compensation of existing lines is typically recommended, over constructing new lines, to boost power transmission capacity and stability margin[1]. On the other hand, fixed series capacitor compensation is a source of SSR phenomenon not only in new power plants but also in existing generating power systems and power electronic equipment[2]. The primary source of the SSR phenomenon is the electrical resonance frequency, which is smaller in amplitude to the natural system frequency. It usually occurs if the natural mechanical frequency of the turbine matches with the electrical resonance frequency—the SSR oscillation results in the slip voltage of SSR frequency at the rotor circuit. The reason for the shaft failure is due to the oscillating electromagnetic torque, which consequently results in the failure of the shaft. This incident invites serious research on the SSR phenomenon in DFIGs and series compensated transmission lines[3].

The SSR phenomena were first discussed in 1937, and it resurfaced in the 1970s after two shaft collapses at the Mohave site received attention[4]. A decade ago, noticed the potential relationship between wind farms and series capacitors. Due to increasing levels of wind power, SSR occurrences have been observed in series-compensated DFIG based wind power plants (WPP) in the south of Texas, the southwest of Minnesota, and the north of China (Hebei province)[5]. Following the incidence, extensive research into the SSR problem among DFIGs and series-compensated power systems was conducted. According to reference [5], DFIGs are more susceptible to SSR in series compensated systems than fixed-speed wind turbine generators and wind generators connected via back-to-back converters.

Several previous research works discuss detailed mathematical modeling of a DFIG-based WPP[3],[6]–[10]. However, the RSC, GSC dynamics have not been considered for the

https://doi.org/10.36375/prepare_u.iei.a221

modeling and system analysis. The RSC and GSC have been considered to be a constant voltage source model[3],[7],[8] and a constant DC connection voltage amongst the RSC and the GSC. The RSC, as well as GSC must be simulated and included in the system modelling for internal dynamic analysis as they can introduce significant impact in the overall system dynamics. The dynamics of DFIG, RSC and GSC can be included in the stability analysis to improve its accuracy.

Now, because this modelling being more accurate, we have included the dynamics of RSC and GSC for overall system modeling and stability analysis via eigenvalue analysis as well as time-domain analysis in this work. The significant information regarding the damping of each mode and its oscillatory frequency can be achieved from the above analysis. The Eigenvalues enables to know the different modes of oscillations available in the complex electrical system

The whole paper is divided into the following sections: Section A discusses the detailed modeling of DFIG based WPP, which is linked with series compensated long-distance transmission line. The result and discussion part are given in section B, in which eigenvalue analysis, as well as time-domain analysis, is carried out. The conclusion is drawn in section C. In the last section, the Appendix shows the system parameters taken.

A. System modeling:

The IEEE first benchmark model has been adapted to analyze the SSR phenomenon in DFIG based wind energy conversion system instead of a synchronous generator integrated with a series compensated transmission line, as shown in Fig. 1[11]. A 161kV series compensated long transmission line connects an infinite bus to a 100MW DFIG based wind power plant (WPP) in this study. The DFIG has a 575 V rated voltage and a 60 Hz frequency. The 67 wind turbines that make up the 100-megawatt wind farm each having a 1.5MW capacity. In the Appendix, parameters for both the aggregated model and the single model are given.

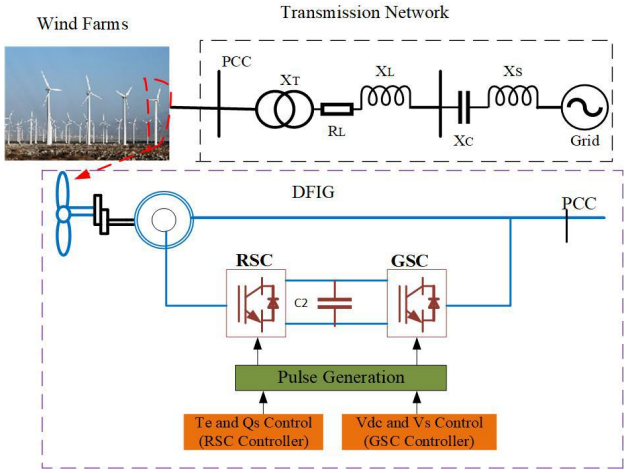


Fig. 1 IEEE first benchmark model of DFIG-based WPP

1) Modeling of wind turbine: A wind turbine's mechanical output torque is proportional to the wind speed and can be expressed as:

$$Tm = \frac{\rho R A C_p V_{wi}^2}{2\lambda} \quad (1)$$

Where ρ represents the air density(kgm^{-3}), R defines wind turbine rotor radius(m), A denotes the blade sweep area(m^2), V_{wi}^2signifies the wind speed(m/s), and C_p denotes the blade power coefficients, which is determined by the blade pitch angle θ and tip speed ratio as fallow:

$$C_p = 0.5\left(\frac{RC_f}{\lambda} - 0.022\theta - 2\right) e^{-0.255\left(\frac{RC_f}{\lambda}\right)} \quad (3)$$

The tip speed ratio is expressed as:

$$\lambda = \frac{\Omega_m R}{V_{wi}} \quad (4)$$

Ω_m denotes the mechanical angular velocity (rad/s)

Fig. 2 depicts the link between mechanical power, rotor speed (pu), and wind speed for a 1.5MW wind turbine when the pitch angle θ is zero with 200ft rotor radius is assumed.

https://doi.org/10.36375/prepare_u.iei.a221

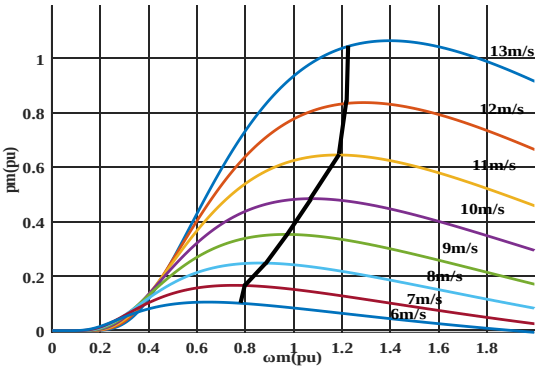


Fig. 2 relationship between wind power wind speed and rotor speed

We can develop a lookup table as shown in Table 1 that lists the various wind speeds and the rotor speed, maximum mechanical power, and shaft torque. It should be emphasized that in the small-signal analysis technique, Table 1 depicts the wind turbine model. This Table, by including a torque control loop, can obtain the desired optimum rotor speed.

TABLE I
Mechanical Power, Torque, Rotor Speed (Pu), and Wind Speed

V_{wi}	7m/s	8m/s	9m/s	10m/s	11m/s	12m/s
ω_m	0.8	0.88	0.97	1.06	1.18	1.2
P_m	0.24	0.28	0.39	0.47	0.64	0.84
T_m	0.3	0.32	0.4	0.46	0.54	0.7

2) Modelling of Shaft System: The two-mass system often represents torsional dynamics in power system analysis. The first mass is represented by wind turbine shaft, which has a low speed, and a DFIG high-speed shaft system represents the second mass system; and A two-mass system's state-space representation is as follows:

$$\frac{d}{dt}\begin{bmatrix}\omega_t\\ \omega_r\\ T_g\end{bmatrix}=\begin{bmatrix}\frac{-D_t-D_{tg}}{2H_t} & \frac{D_{tg}}{2H_t} & \frac{-1}{2H_t}\\ \frac{D_{tg}}{2H_g} & \frac{-D_g-D_{tg}}{2H_g} & \frac{1}{2H_g}\\ K_{tg}\omega_g & -K_{tg}\omega_e & 0\end{bmatrix}\begin{bmatrix}\omega_t\\ \omega_r\\ T_g\end{bmatrix}+\begin{bmatrix}\frac{T_m}{2H_t}\\ \frac{-T_e}{2H_t}\\ 0\end{bmatrix} \quad (5)$$

The wind turbine, DFIG rotor speed, and internal torque of the model are represented by state variablesω_t, ω_r and T_g , respectively.T_m and T_e are the mechanical and electromagnetic torque of the generator, respectively, and these are shaft system models of the input side. D_t and H_t are the damping coefficient and inertia constant of the wind turbine, respectively, and H_g is the inertia constant of the wind generator. D_{tg} and K_{tg} are the damping coefficient and shaft stiffness of the coupling between two masses; these are the constant of a two-mass model system. The optimal wind turbine torque T_m is obtained from table1 at any required wind speed, whereas T_e it could be written as air gap flux leakage and currents of the DFIG is express:

$$T_e=\Psi_{qm}i_{dr}-\Psi_{dm}i_{qr} \quad (6)$$

$$\Psi_{qm}=X_m(i_{qs}+i_{qr}) \quad (7)$$

$$\Psi_{dm}=X_m(i_{ds}+i_{dr}) \quad (8)$$

3) Modelling of DFIG

The modelling of DFIG with rotor side converter, a dynamic model of the sixth order is used. The model is characterized as**:**

$$\dot{x}=A_DX_D+B_DU_D \quad (9)$$

where the state variables are denoted by X_D, U_D denotes the inputs of DFIG, B_D denotes the constants of electrical parameter, and A_D Represents the constants of state variables.

$$X_D=[i_{qs};\ i_{ds};\ i_{0s};\ i_{qr};i_{dr};i_{0r};] \quad (10)$$

$$U_D=[v_{qs};v_{ds};v_{0s};v_{qr};v_{dr};v_{0r};] \quad (11)$$

The subscripts s, r, d, q, and 0 signify the stator and rotor sides, respectively, as well as direct, quadrature, and zero sequences. The parameters i0s, i0r, v0s, and v0r are equal to zero because the system is balanced.

$A_D$and $B$is expressed as:

$$A_D=-B_D\begin{bmatrix}R_S & \frac{\omega_s}{\omega_b}X_s & 0 & 0 & \frac{\omega_s}{\omega_b}X_m & 0\\ -\frac{\omega_s}{\omega_b}X_s & R_S & 0 & -\frac{\omega_s}{\omega_b}X_m & 0 & 0\\ 0 & 0 & R_S & 0 & 0 & 0\\ 0 & \frac{\omega_s-\omega_r}{\omega_b}X_m & 0 & R_r & \frac{\omega_s-\omega_r}{\omega_b}X_r & 0\\ -\frac{\omega_s-\omega_r}{\omega_b}X_m & 0 & 0 & -\frac{\omega_s-\omega_r}{\omega_b}X_r & R_r & 0\\ 0 & 0 & 0 & 0 & 0 & R_r\end{bmatrix}$$

$$B_D=\omega_b\begin{bmatrix}X_s & 0 & 0 & X_m & 0 & 0\\ 0 & X_s & 0 & 0 & X_m & 0\\ 0 & 0 & X_{ls} & 0 & 0 & 0\\ X_m & 0 & 0 & X_r & 0 & 0\\ 0 & X_m & 0 & 0 & X_r & 0\\ 0 & 0 & 0 & 0 & 0 & X_{lr}\end{bmatrix}^{-1} \quad (12)$$

Except for ω_b, ω_r, and ω_s, which are in rad/sec, all parameters and variables are in per-unit.

https://doi.org/10.36375/prepare_u.iei.a221

4) Modelling of DC link: A first-order model describes the dynamics of the capacitor in the dc link between the GSC and the RSC.

$$C_{dc}V_{dc}\frac{dV_{dc}}{dt} = P_r - P_g \quad (13)$$

$$P_r = \frac{1}{2}\left(v_{qr}i_{qr} + v_{dr}i_{dr}\right) \quad (14)$$

$$P_g = \frac{1}{2}\left(v_{qg}i_{qg} + v_{dg}i_{dg}\right) \quad (15)$$

The RSC and GSC's active powers are P_r and P_g, respectively. V_{dc} is the actual value whereas, $v_{qr}, i_{qr}, v_{dr}, i_{dr}, v_{qg}, i_{qg}, v_{dg}, i_{dg}$ are in per-unit.

5) Series compensated transmission line Modeling: Generally, the dynamics of the series compensated model are ignored, but SSR analysis is essential. The transmission line can be electrically depicted as lumped series RLC circuit and adopts the synchronous reference frame to modeling the series compensated transmission line. Per phase dynamics equations is expressed as:

$$R_l.i_l + L\frac{di_l}{dt} + v_c = v_s - V_B \quad (16)$$

$$C\frac{dv_c}{dt} = i_l \quad (17)$$

The compensated transmission's state-space representation is expressed as:

$$\frac{d}{dt}\begin{bmatrix} i_{ql} \\ i_{dl} \\ v_{cq} \\ v_{cd} \end{bmatrix} = \begin{bmatrix} -\frac{R_l}{X_l} & -\omega_e & -\frac{1}{X_l} & 0 \\ -\omega_e & -\frac{R_l}{X_l} & 0 & -\frac{1}{X_l} \\ X_c & 0 & 0 & -\omega_e \\ 0 & X_c & \omega_e & 0 \end{bmatrix}\begin{bmatrix} i_{ql} \\ i_{dl} \\ v_{cq} \\ v_{cd} \end{bmatrix} + \omega_b \begin{bmatrix} \frac{v_{qs}-V_{Bq}}{X_l} \\ \frac{v_{ds}-V_{Bd}}{X_l} \\ 0 \\ 0 \end{bmatrix} \quad (18)$$

Where i_{dl} i_{ql} is the transmission line's direct and quadrature axis current, respectively, and v_{cd} are the quadrature also direct axis voltage across the capacitor. The Voltage's quadrature and direct axes of the infinite are V_{Bq} and V_{Bd}, respectively. The synchronous reference speed and the base speed are both 1pu or 377rad/s.

6) 6) Modeling of DFIG converter: This study aims to look at both RSC and GSC converter controllers. Vector-controlled techniques is adopted in this study. The control loop is shown in Fig.2 and Fig.3. For the RSC controller, Table1 is used for the obtain the optimal reference torque. When the wind speed exceeds the rated speed, the value remains constant. When the wind speed is less than the rated speed, the reference torque is the torque that provides the maximum rotational speed at the measured rotational speed. Using this lookup table, the wind turbine can extract the maximum amount of wind energy feasible.

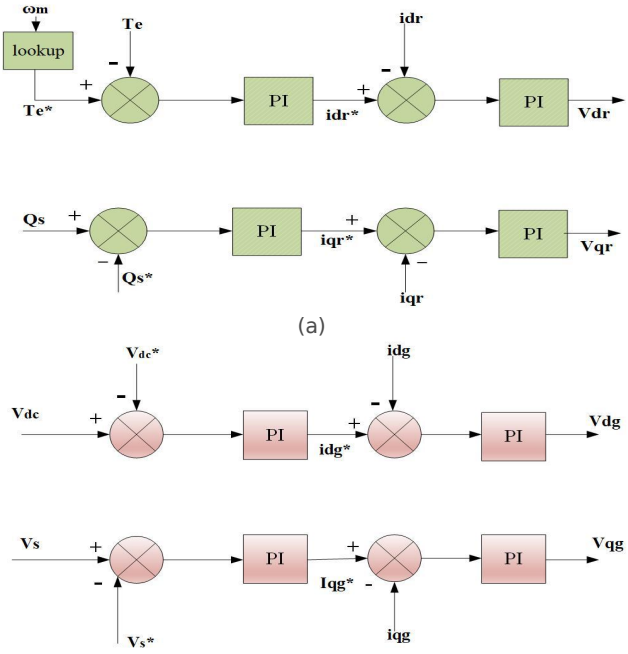


Fig. 3 (a) RSC, (b) GSC control loop

7) Integrated system modelling: To understand the complete system, it is useful to think of it as an integration of the various subsystems. Equation expresses the entire system model (25).

$$[\dot{X}] = [A_{sys}][X_{sys}] + [B_{sys}][U_{sys}] \quad (19)$$

Where $[X_{sys}] = [X_{shaft}\, X_{DFIG}\, X_{tl}\, X_{dc}]$ (20)

The system matrix of the entire system is given by

$$[A_{sys}] = \begin{bmatrix} A_{shaft} & 0 & 0 & 0 \\ 0 & A_{DFIG} & 0 & 0 \\ 0 & 0 & A_{tl} & 0 \\ 0 & 0 & 0 & A_{dc} \end{bmatrix} \quad (21)$$

These interconnected subsystems are able to communicate with each other via state and control variables in the individual subsystems. Algebraic Equations of v_{ds}, v_{qs}, i_{ds} and i_{qs} are used to link the controllers. Now v_{ds}, v_{qs}, i_{ds} and i_{qs} expressed as:

$$v_{ds} = v_{dg} + X_{tg}i_{qg} \quad (22)$$

$$v_{qs} = v_{qg} + X_{tg}i_{dg} \quad (23)$$

$$i_{ds} = i_{dg} - i_{dl} \quad (24)$$

https://doi.org/10.36375/prepare_u.iei.a221

$i_{qs}=i_{qg}-i_{ql}$
(25)
Developed torque and stator reactive power is given by:
$T_g = X_m\left(i_{dr}i_{qs}-i_{ds}i_{qr}\right)$
(26)
$Q_s = \frac{1}{2}\left(v_{qs}i_{ds}-v_{ds}i_{qs}\right)$
(27)
A state-space model is used to investigate the system's small signal stability, which leads to time-domain simulation analysis and eigenvalue computation.

B. Result and discussion

A) *Eigen value analysis:* The eigenvalue approach is one of the most critical approaches for studying a dynamic system's small-signal stability. Based on the order of the system matrix, this analysis generates a collection of eigenvalues. The eigen values can be used to determine a system's stability; if the real component of an eigen value is positive, the system is unstable; if the real part is negative, the system is stable. The frequency of oscillations is represented by the imaginary portion. The size of the system matrix (A_{sys}) is determined in Section A, which is the 22x22 size of the matrix, resulting in 22 eigenvalues. By varying the wind speed and compensation level of a transmission line, different eigen values can be determined. Table II displays the set of eigenvalues (from λ_1 to λ_{22}) for fixed wind speeds at 7m/sand at a constant capacitive compensation level of 50%. As shown in Table II, the eigenvalues λ5 and λ6 are found unstable at these operating circumstances, and the system becomes increasingly unstable.

The remaining computed eigenvalues operate in a variety of ways. Electromechanical modes with eigen values of λ_7,-λ_{12} are reported to be responsible for electromechanical system dynamics. λ_1,λ_2, and λ_3,λ_4 belong to the extremely high frequency and super synchronous modes, respectively, because the frequency of these dynamics is higher than the system's, they are always stable. The λ_{13},λ_{22} do not have any frequency terms and are classified as oscillatory as positive modes or non-oscillatory as a negative mode, with amplitudes which are variable as a function of compensation levels as well as wind speeds. As compensation increases, the real part of λ_5,λ_6 grows in size, and the system becomes increasingly unstable, as seen in Table III. The system stability is observed to be dependent on the percentage of compensation offered to the transmission line based on the eigenvalues. The system gets more stable as the wind speed increases. The time-domain simulation of this system, as well as the RSC and GSC controllers, are detailed in the next section.

TABLE II
EIGEN VALUE ANALYSIS OF THE COMPLETE SYSTEM MODEL WITH 50% COMPENSATION AND A WIND SPEED OF 7 M/S

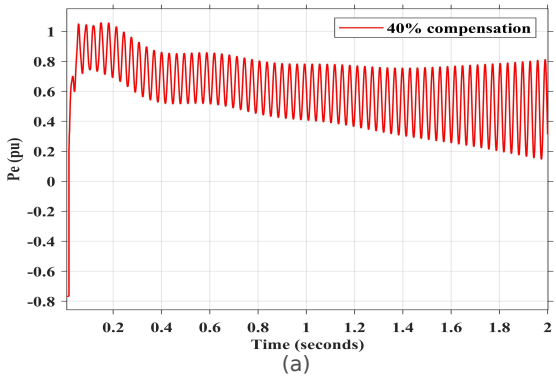
Mode	Eigenvalue	Frequency (Hz)	Description
$\lambda_{1,2}$	$-1790\pm j889$	141	Very high freq.
$\lambda_{3,4}$	$-19.2\pm j568.2$	90.4	Sup-sync
$\lambda_{5,6}$	$\mathbf{1.6\pm j187.8}$	**29.9**	**SSR**
$\lambda_{7,8}$	$-52.4\pm j7.9$	1.25	Electromech.
$\lambda_{9,10}$	$-0.63\pm j18.8$	2.86	Electromech.
$\lambda_{11,12}$	$-6.41\pm j8.6$	1.36	Electromech.
$\lambda_{20,21}$	$-0.0\pm j0.0$	0	PI Controller
λ_{13}	5.9	0	No oscillatory
λ_{14}	-7.4	0	No oscillatory
λ_{15}	-5.8	0	No oscillatory
λ_{16}	-5.6	0	No oscillatory
λ_{17}	1.9	0	No oscillatory
λ_{18}	-48.2	0	No oscillatory
λ_{19}	-37.7	0	No oscillatory
λ_{22}	-0.0	0	No oscillatory

TABLE III
EIGEN VALUE ANALYSIS OF MODE 3($\lambda_{5,6}$) AT VARYING WIND SPEED AND % OF SERIES COMPENSATION

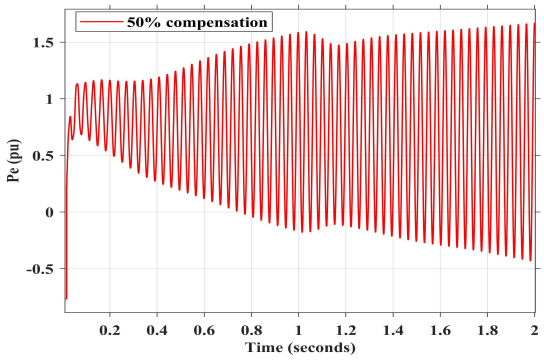
%Compensation	Wind Speed	Mode3 ($\lambda_{5,6}$)	
		Eigen value	Frequency (Hz)
40	7m/s	$\mathbf{0.01\pm j207}$	32.96
	9m/s	$0.04\pm j206.9$	32.94
	11m/s	$0.12\pm j206.8$	32.92
50	7m/s	$\mathbf{1.6\pm j187.8}$	29.90
	9m/s	$\mathbf{1.5\pm j187.7}$	29.88
	11m/s	$\mathbf{1.4\pm j187.5}$	29.85
60	7m/s	$\mathbf{3.2\pm j170.9}$	27.21
	9m/s	$\mathbf{3.1\pm j170.7}$	27.18
	11m/s	$\mathbf{2.9\pm j170.4}$	27.13

https://doi.org/10.36375/prepare_u.iei.a221

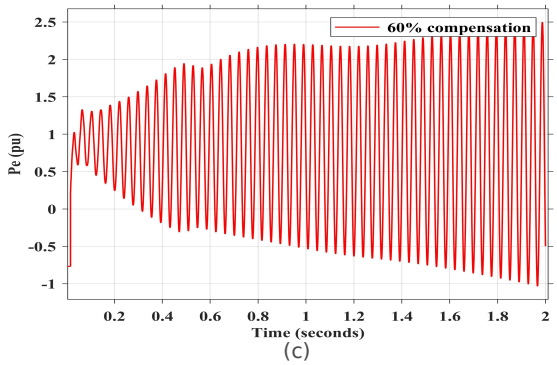
B) Time-domain analysis: MATLAB/Simulink has been used for the time-domain analysis to verify the results of eigenvalue analysis. At a constant wind speed (7 m/s), Fig. 4 displays the Pe electrical dynamic response with various compensation values. The dynamic reaction is initiated when the compensation level of 40% is increased to 50% and 60%, respectively, at time t=2s. At a specific wind speed (9 m/s), Fig. 5 displays the Pe electrical dynamic response with various compensation values. As illustrated in Fig.4 and 5, The damping decreases as the compensating level rises. The result reveals that as the compensation level is raised, the SSR frequency rises while the damping decreases. As seen in Fig. 5, the SSR damping improves with increasing wind speed.



(a)

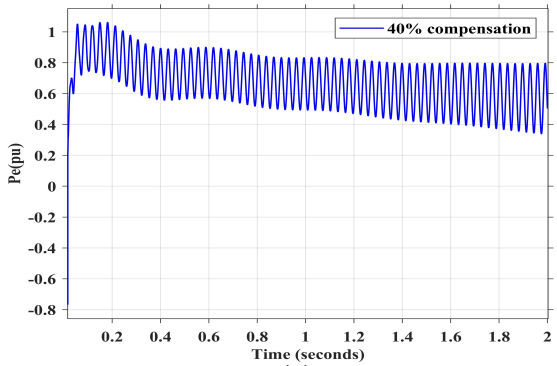


(b)

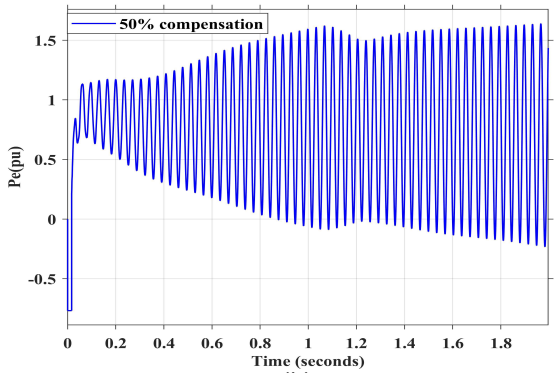


(c)

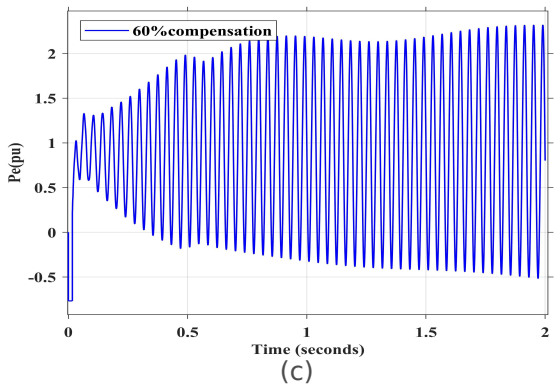
Fig. 4 Electrical power dynamics response at (a)40%, (b) 50% and (c) 60% compensation at fixed wind speed 7m/s



(a)



(b)



(c)

https://doi.org/10.36375/prepare_u.iei.a221

Fig. 5 Electrical power dynamics response at (a)40%, (b) 50% and (c) 60% compensation at fixed wind speed 9m/s

As a result, the system becomes unstable at a wind speed of 7 m/s due to the negative damping of the oscillation. The system is stable at greater wind speeds. These findings are consistent with the notion that the dominant mode becomes more unstable when wind speed falls.

C. Conclusion

For the small-signal stability investigation, a detailed mathematical model of the 100MW grid-connected DFIG based WPP based on the IEEE first benchmark model, as well as a dynamic model of the RSC and GSC converters, has been done. The numerous modes of oscillations present in the integrated system can be easily discovered using eigenvalue analysis. The stability of the various oscillating modes can be identified by analysing the nature of eigenvalues. In MATLAB/Simulink, oscillations present in the electromagnetic torque produced for various compensation levels are investigated, and the eigenvalue analysis is validated. In this paper, SSR analysis has been carried out using an eigenvalue technique for various wind speeds and capacitance compensation levels and got different oscillation modes, and the above result has been validated using time-domain analysis. The fallowing conclusion has obtained:

- The real component magnitude of the eigenvalue increases as the fraction of series compensation is raised, and the system becomes increasingly unstable.
- The increase in wind speed velocity decreases the amplitude of the real part of eigen value, resulting in a better stable system.

APPENDIX

TABLE IV
DFIG PARAMETERS

Rated power	1.5MW	100MW
Rated Voltage	575V	575V
Rated Frequency	60Hz	60Hz
R_S	0.023	0.023
X_{lS}	0.18	0.18
R_r	0.016	0.016
X_{lr}	0.16	0.16
X_M	2.9	2.9
DC-link voltage	1200V	1200V
DC-link capacitor	10000μF	67*10000μF

TABLE V
TRANSMISSION LINE PARAMETERS

Transformer Ratio	575V/161kV
Base power	100MVA
R_l	0.02
X_l	0.5
X_C at 50% compensation level	64.8 ohm
X_T	0.14
X_{sys}	0.06

TABLE IV
DRIVE TRAIN PARAMETERS

H_t	4.32 sec
H_g	0.685 sec
D_t	0
D_g	0
D_{tg}	1.5 p. u
K_{tg}	1.11

TABLE V
PI PARAMETERS OF DFIG CONVERTER

K_{P1}	0.2	K_{i1}	0.4
K_{P2}	4	K_{i2}	100
K_{P3}	0.01	K_{i3}	0.04
K_{P4}	4	K_{i4}	100
K_{P5}	0.1	K_{i5}	2
K_{P6}	0.08	K_{i6}	0.5
K_{P7}	0.1	K_{i7}	2
K_{P8}	0.08	K_{i8}	0.5

REFERENCES

[1] M. Khayyatzadeh and R. Kazemzadeh, "Sub-synchronous resonance damping using high penetration PV plant," *Mech. Syst. Signal Process.*, vol. 84, pp. 431–444, 2017.

[2] G. Li *et al.*, "Analysis and Mitigation of Subsynchronous Resonance in Series-Compensated Grid-Connected System Controlled by a Virtual Synchronous Generator," *IEEE Trans. Power Electron.*, vol. 35, no. 10, pp. 11096–11107, 2020.

[3] L. Fan, R. Kavasseri, Z. L. Miao, and C. Zhu, "Modeling of DFIG-based wind farms for SSR analysis," *IEEE Trans. Power Deliv.*, vol. 25, no. 4, pp. 2073–2082, 2010.

[4] A. Adrees, *Risk Based Assessment of Subsynchronous Resonance in AC/DC Systems*. 2017.

[5] J. Shair, X. Xie, L. Wang, W. Liu, J. He, and H. Liu, "Overview of emerging subsynchronous oscillations in practical wind power systems," *Renew. Sustain. Energy Rev.*, vol. 99, no. September 2018, pp. 159–168, 2019.

[6] J. B. Ekanayake, L. Holdsworth, X. G. Wu, and N. Jenkins, "Dynamic modeling of doubly fed induction generator wind turbines," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 803–809, 2003.

[7] Y. Lei, A. Mullane, G. Lightbody, and R. Yacamini, "Modeling of the wind turbine with a doubly fed induction generator for grid integration studies," *IEEE Trans. Energy Convers.*, vol. 21, no. 1, pp. 257–264, 2006.

[8] H. A. Mohammadpour, A. Ghaderi, and E. Santi, "Analysis of sub-synchronous resonance in doubly-fed

https://doi.org/10.36375/prepare_u.iei.a221

induction generator-based wind farms interfaced with gate – controlled series capacitor," *IET Gener. Transm. Distrib.*, vol. 8, no. May, pp. 1998–2011, 2014.

[9] A. S. Trevisan *et al.*, "Analysis of low frequency interactions of DFIG wind turbine systems in series compensated grids," *Electr. Power Syst. Res.*, vol. 191, no. February 2020, p. 106845, 2021.

[10] J. Shair, X. Xie, W. Liu, X. Li, and H. Li, "Modeling and stability analysis methods for investigating subsynchronous control interaction in large-scale wind power systems," *Renew. Sustain. Energy Rev.*, vol. 135, no. August 2020, p. 110420, 2021.

[11] P. S. E. C. IEEE, "IEEE Subsynchronous Resonance Task Force of the Dynamic System Performance Working Group Power System Engineering," *IEEE Trans. Power Appar. Syst.*, vol. PAS-96, no. 5, pp. 1565–1572, 1977.

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/TVT.2021.3131504, IEEE Transactions on Vehicular Technology

Statistical Characterization of the Angle of Arrival in an Underwater Channel due to scattering from the sea surface modeled as Sum of Sinusoids

Manishika Rawat, *Student Member*, IEEE, Brejesh Lall, *Member*, IEEE and Seshan Srirangarajan, *Member*, IEEE

***Abstract*—The objective of this work is the statistical characterization of the angle of arrival (AoA) at the receiver due to scattering from a random sea surface. The sea surface is represented using the sum of sinusoids (SoS) and modeled using the JOint North Sea WAve Project (JONSWAP) sea wave spectrum. The random behavior of potential scatterers along the sea surface, surface wave height, and their impact on the distribution of AoA is evaluated. For coherent scattering, initially, the scatterer distribution is computed for a single realization of the SoS, i.e., a sinusoidal surface. The result obtained is averaged to evaluate the characteristics of the ensemble-averaged SoS surface. For incoherent scattering, the acute randomness of the sea surface leads to every point along the SoS surface acting as a potential scatterer. The overall scattering for any given channel state would be an appropriate combination of coherent and incoherent components. A similar effect manifests itself in the distribution of the overall AoA. Further, the impact of factors such as acoustic signal frequency, wind speed, and the underwater (UW) channel geometry on the parameters of the overall AoA has been studied. The AoA model proposed in this work is general and widely applicable.**

***Index Terms*—Underwater acoustic communication, angle of arrival distribution, coherent scattering, incoherent scattering, sea surface roughness.**

I. Introduction

Underwater (UW) is a time-varying, multipath, and frequency-selective channel where the communication depends on the various external sources such as wind speed above the sea surface, the fetch limit range, temperature of the sea surface, the gravitational force, etc. The most critical issue that UW acoustic communication face is the lack of an appropriate model representing the UW channel [1]. Due to the blowing wind, the sea surface is always moving, and the channel changes with space. It has been observed that the sea surface waves that cause the scattering of incident acoustic signals are the most prominent factors that make the UW channel complex and time-variable [2]. Thus, UW acoustic communication is a highly random process, and therefore a statistical analysis is warranted.

Sea surface modeling plays a vital role in the characterization of an UW channel [3]. For a given wind speed above the sea surface, a certain amount of energy is transferred to give rise to random fluctuations in the sea surface. For a given wind speed, sea waves comprise of several frequencies derived from the wave spectrum ([4], ch. 2). Sea wave spectrum provides an insight into the distribution of wave energy as a function of wave frequency. Various spectrum models have been introduced in the literature for better representation of surface waves. The two most popular spectrum models used are Pierson-Moskowitz (PM) spectrum [5] and JOint North Sea WAve Project (JONSWAP) spectrum [6]. In the PM model, the wave spectral density is a function of wind speed and wave frequency, while the JONSWAP model is an extension of the PM model for fetch limited regions. In this work, we represent the sea surface using sum of sinusoids (SoS) with the frequencies generated from the JONSWAP spectrum model, as proposed in [7]. This model gives the sea surface distribution as a function of wind speed.

In UW acoustic communication, an acoustic signal reaches the receiver (Rx) either through the line of sight or after undergoing scattering from the sea surface and bottom. This work focuses on scattering from a fluctuating sea surface. Scattering may be coherent or incoherent, depending on the roughness of the surface, which is quantified by the value of the Rayleigh parameter, $\chi = 2h\Lambda \sin\theta_g$ [8], [9]. Here, Λ is the acoustic wave number determined by $\Lambda = 2\pi f/c$ (m^{-1}), where c is the speed of sound in (m/s) in an UW channel, f is the acoustic signal frequency in (Hz), h is the root mean square (RMS) sea surface wave height in (m), and θ_g is the grazing angle associated with the nominal specular path from the mean sea level. Coherent scattering dominates when the Rayleigh parameter is small, i.e., $\chi < 1$ or when the sea surface is acoustically smooth. The scattered signal is in phase, and thus the acoustic signal gets reflected by the sea surface and bottom. However, as the surface roughness increases, the scattering pattern at the surface changes. At relatively higher frequencies and wind speeds, i.e., $\chi >> 1$, every point on the surface acts as a potential scatterer. This phenomenon is called incoherent scattering, where the phase of the scattered signal is highly random. This results in an omnidirectional scattering pattern around the scattering point on sea surface ([10], ch. 3). Experimental measurements such as those presented in [6], [11] indicate that at frequencies above 10 kHz, the RMS wave height of most naturally occurring sea surfaces results in the Rayleigh parameter being much greater than 1. The implication is that high-frequency signals undergo incoherent scattering from the sea surface on an ensemble mean basis. For intermediate values of the Rayleigh parameter, the scattering should be an appropriate combination of coherent and incoherent components depending on the sea surface roughness.

In a radio frequency channel, due to a large number of scatterers, the signal reaches the Rx from all directions isotropically [12]. In contrast, in shallow water communication, the acoustic signal gets scattered from the sea surface and bottom

0018-9545 (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 December 06,2021 at 06:38:40 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/TVT.2021.3131504, IEEE Transactions on Vehicular Technology

in a restricted geometry before reaching the Rx, resulting in a specific distribution of the angle of arrival (AoA). Therefore, there is a direct dependence of the path gain statistics on the AoA. The directional communication offered by an UW channel due to the constrained geometry makes AoA a vital characteristic of an UW channel. Applications such as pollution monitoring and seismic exploration require communication between moving platforms. Since the platforms are mobile, it is preferable to use directional receivers. For such systems, the AoA based channel modeling offers an advantage in terms of modeling accuracy. AoA analysis is essential for applications such as source localization and beamforming [13]–[18]. Besides, various channel characteristics for mobile transmitter (Tx)/Rx such as autocorrelation function (ACF), Doppler spread, power azimuth spectrum depend on the AoA and its distribution [19]. AoA based detection of UW acoustic Orthogonal frequency-division multiplexing (OFDM) signal was proposed in [20]. Despite all its potential, limited work has been done to evaluate the AoA at the Rx in an UW channel.

In [21], authors have presented the plots of the AoA distribution at the Rx for a path scattered from the sea surface modeled using the JONSWAP spectrum. However, this is empirical modeling, and the expressions for the AoA distribution are not derived. The AoA distribution for specularly reflected paths (coherent scattering) from a smooth sea surface and bottom in an UW channel has been derived in [19]. The derived expressions for the AoA have been used to evaluate the ACF and Doppler spectrum for relative Tx/Rx motion. However, the assumption of the smooth surface limits the application of this analysis. In [9], [22], authors have evaluated the AoA distribution due to incoherent scattering from the flat sea surface for single, double, and multi-bounce paths. In [22], authors have proved that the AoA distribution for incoherent scattering from a flat surface matches reasonably well with the distribution obtained for the sinusoidal surface at low wind speeds. This is because, at low wind speeds, the sea surface almost appears flat for a certain depth of the Tx or Rx. However, a general distribution for AoA from a fluctuating sea surface has never been evaluated and deserves a complete study due to its enormous research potential and application. Our work is an attempt to bridge this gap to an extent.

The major contribution of this work to the existing state of the art can be summed up as follows:

(i) In this work, we derive the statistical distribution of the AoA at the Rx for coherent and incoherent scattering from a fluctuating sea surface. The surface is represented using SoS and modeled using the JONSWAP spectrum. Since UW is a time-varying channel, the distribution is derived for an ensemble-averaged sea surface obtained after averaging over several realizations of the sinusoidal surface.

(ii) A general expression for the overall AoA at the Rx is proposed, and its distribution is evaluated. The AoA model proposed can be applied to any UW channel with Gaussian sea surface roughness.

(iii) The impact of various factors such as UW channel geometry (propagation range, Tx/Rx depth), acoustic signal frequency, and wind speed above the sea surface on the parameters of the overall AoA is examined. This analysis can be used to achieve the desired performance by trading off acoustic signal frequency and channel geometry for given wind speed.

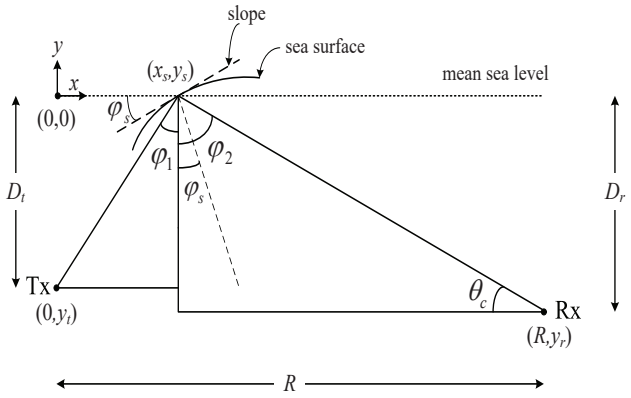


Fig. 1. Geometry of an UW channel depicting scattering from a part of a fluctuating sea surface.

The primary motivation behind this work is to provide a fundamental and general model for AoA at the Rx in an UW channel, which is applicable for different sea surface conditions and geometries. We have tried to incorporate the randomness of the sea surface and evaluate its impact on the AoA under different environmental and geometrical conditions. This modeling of AoA at the Rx brings us closer to the practical scenario.

The paper has been organized as follows. Section II presents the description of the system model and the sea surface representation using SoS. Section III consists of the derivation of the distribution of scatterers, the surface wave height, and the AoA assuming coherent scattering. On a similar note, Section IV contains the derivation of the distribution of scatterers, the surface wave height, and AoA assuming incoherent scattering. Section V includes the expression and derivation of overall AoA at the Rx for any given channel condition. In Section VI, we present the results and discussions. We finally conclude in Section VII.

II. System Model

We assume a two-dimensional cross-section of an UW channel and one-dimensional sea surface, as shown in Fig. 1. The Tx and Rx are assumed to be omnidirectional such that they can ensonify the entire scattering region in between [23]. The UW channel is assumed to be an isovelocity channel such that the acoustic signal travels in a straight line. In this work, we deal with small depths of Tx/Rx (few meters) to evaluate the impact of surface fluctuations (with an amplitude of fractions of a meter) on the AoA. For such small depths, the error introduced due to isovelocity approximation would be minimal. The acoustic signal gets scattered from the sea surface before reaching the Rx, and we consider only forward scattering. The scattering may be coherent or incoherent, depending on the sea surface roughness. The purpose of this

0018-9545 (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 December 06,2021 at 06:38:40 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/TVT.2021.3131504, IEEE Transactions on Vehicular Technology

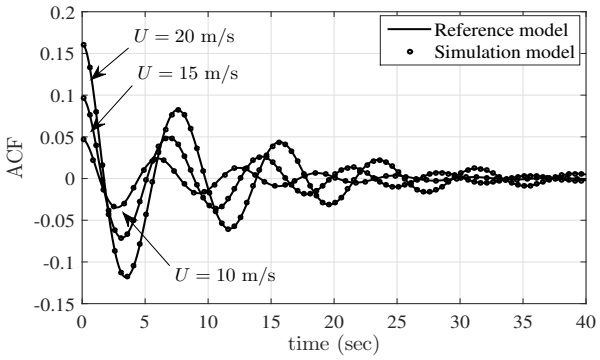


Fig. 2. Autocorrelation function of y obtained using the reference model (JONSWAP spectrum model) and the simulation model (SoS) for different values of wind speed at $N = 40$.

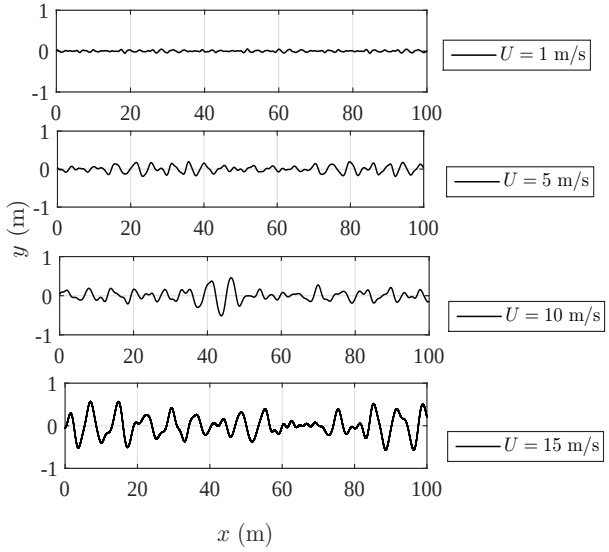


Fig. 3. Sea surface profile along the range for a surface represented using SoS and modeled using JONSWAP spectrum for different values of wind speed and $F = 100$ km.

work is to determine the AoA distribution at the Rx for a random sea surface and evaluate the impact of various surface conditions on the AoA. We represent the sea surface using the SoS and model it using JONSWAP spectrum. JONSWAP spectrum modeling brings the sea surface close to the practical scenario.

A. JONSWAP Sea Wave spectrum

The JONSWAP frequency spectrum is an empirical expression for the spectral density of ocean displacements due to wind-blown waves acting over a finite fetch [24]. The wave spectrum so obtained is a function of wind speed, wave frequency, and fetch limit range. The symmetrical wave spectrum can be expressed as [7]

$$S_{yy}(\omega) = \frac{\alpha_J g^2}{2|\omega|^5} \exp\left(-1.25\left(\frac{\omega_p}{\omega}\right)^4\right) \gamma^{e^{-\frac{(|\omega|-\omega_p)^2}{2\sigma_J^2\omega_p^2}}}, \quad -\infty < \omega < \infty. \tag{1}$$

Here, $\alpha_J = 0.076(U^2/Fg)^{0.22}$, where F is the fetch defined as the distance over which the wind blows with constant velocity, and $\omega_p = 22(g^2/FU)^{1/3}$ is the peak wave frequency. The parameter γ is the peak enhancement factor given by $\gamma = 3.3$, $\sigma_J = 0.07$ for $|\omega| \leq \omega_p$, and $\sigma_J = 0.09$ for $|\omega| > \omega_p$, $g = 9.81\ m/s^2$ is the gravitational acceleration, and ω is sea wave frequency in rad/sec. U is the speed of the wind blowing at 10 m above the sea level. The RMS wave height for a JONSWAP spectrum is given by [7]

$$h = \int_0^\infty S_{yy}(\omega)d\omega = \sqrt{1.67 \times 10^{-7}\frac{U^2F}{g}} = \sqrt{m_0}. \tag{2}$$

Here, the zeroth moment m_0 is the area under the curve.

B. Representing the sea surface as SoS

From the concept of the wave spectrum, we observe that sea wave at any point of time constitutes a range of frequency components. These frequencies are a function of wind speed, as a wind of specific energy can only create a particular kind of wave. Thus, for constant wind speed, the energy of the sea surface remains unchanged [4]. Based on the above observations, we represent the sea surface using SoS as

$$y(x,t) = \sum_{n=1}^{N} a_n \sin(\omega_n t + k_n x + \phi_n). \tag{3}$$

Here, y is the sea surface wave height, and x is the range along the two-dimensional system model, as shown in Fig. 1. N is the number of frequencies that constitute a sea wave. a_n and ω_n represent the amplitude (in m) and frequency (in rad/s) of the nth wave respectively. k_n is the wavenumber given by $k_n = \omega_n/v_s$, where v_s (in m/s) is the velocity at which sea surface drifts or moves. ϕ_n represents the nth realization of ϕ, where ϕ is modeled as an i.i.d. random variable, which is uniformly distributed in the range $(0, 2\pi]$. The approximate value of a_n and ω_n for a sea surface modeled using JONSWAP spectrum can be computed by using the method of equal areas (MEA) given in [7].

In order to verify if the sea surface representation given in (3) upholds the JONSWAP spectrum, we have plotted the autocorrelation function (ACF) of sea surface wave height in Fig. 2. ACF has been obtained using the reference model (eq (4) of [7]) and the simulation model (eq (13) of [7]) for different values of wind speed. The reference model has been obtained using the JONSWAP spectrum, while the simulation model is obtained using the SoS sea surface representation. The exact match of the simulation model with the reference model for different wind speeds validates the SoS representation. At $N = 40$, the ACF of the proposed SoS model matches precisely with the JONSWAP spectrum. Thus, we work with a fixed value of $N = 40$ throughout the analysis.

Fig. 3 gives a pictorial illustration of the sea surface profile at different wind speeds, generated using MEA. We observe that the sea surface wave height and correlation length increases with wind speed, which is also evident from Fig. 2.

0018-9545 (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 December 06,2021 at 06:38:40 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/TVT.2021.3131504, IEEE Transactions on Vehicular Technology

Algorithm 1: To determine the coordinates of the scatterers corresponding to specular paths

initialization a_n, ω_n, $x \subset [0, R]$, $\mu = [\,]$
for *each value of the resolution of* x **do**
  $y_n = a_n \sin(k_n x + \phi'_n)$ sinusoidal surface
  $m = a_n k_n \cos(k_n x + \phi'_n)$ slope of the sea surface at a given value of x
  $\varphi_s = \tan^{-1} m$ angle corresponding to slope m
  $\varphi_1 = \tan^{-1}\left(\frac{x_s}{y_t}\right)$
  $\varphi_2 = \tan^{-1}\left(\frac{R-x_s}{y_s-y_r}\right)$
  $\varphi_{\text{inc}} = \varphi_1 + \varphi_s$ incident angle
  $\varphi_{\text{sc}} = \varphi_2 - \varphi_s$ scattered angle
  if $\varphi_{\text{inc}} = \varphi_{\text{sc}}$ **then**
    $\mu = [\mu\ x]$ scatterer coordinates
  end
end

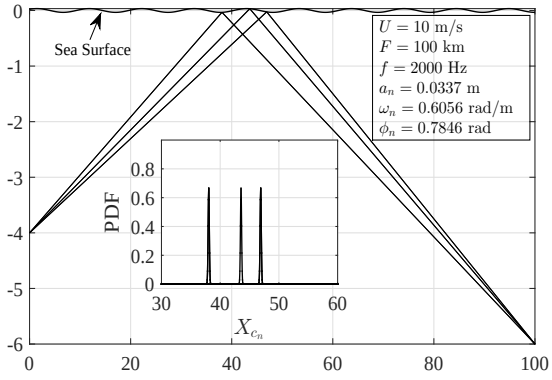


Fig. 4. Figure depicting the occurrence of specular paths derived using algorithm 1 for a realization of sea surface at $R = 100$ m, $D_t = 4$ m, and $D_r = 6$ m. At inset is the PDF of X_{c_n} using (5).

This is because, as wind speed increases, the surface fluctuations increase. The surface becomes more random, which reduces the correlation factor. This affects the scattering, and hence the AoA distribution at the Rx. This has been discussed in detail in the following sections.

III. AoA distribution at the receiver for coherent scattering

In this section, we aim to assess the randomness of potential scatterers (point of interaction of acoustic signal with sea surface), randomness induced in surface wave height, and the distribution of AoA at the Rx, for coherent scattering. In coherent scattering, the acoustic signal gets reflected by the sea surface and bottom due to small surface roughness. We start by analyzing the scattering behavior for a single realization of the sea surface, i.e., a sinusoid. The results obtained for a sinusoidal surface can be averaged to evaluate the characteristics of the ensemble-averaged SoS surface. The n-th realization of the sea surface can be expressed as

$$y_n(x,t) = a_n \sin(k_n x + \phi'_n), \tag{4}$$

where, $0 \leq x \leq R$ and ϕ'_n is the n-th realization of ϕ', where $\phi' \sim \mathcal{U}[\Omega t, 2\pi + \Omega t]$ for a given observation time 't'. Here, $\Omega = [\omega_1, \omega_2, \ldots, \omega_N]$. The coordinates of the scatterers on the sea surface can be determined by the algorithm 1. In this algorithm, the scatterer coordinates are computed using an exhaustive search across the length of the sea surface. The incident and scattered angle for a path scattered from a surface of slope m can be computed using the geometry given in Fig. 1. For each value of the resolution of x, the specular path would exist if the incident angle is equal to the scattered angle. The corresponding values of $x = \mu$ are the required scatterer coordinates. Fig. 1 gives the geometry of an UW channel depicting the evaluation of a specular path for a curved sea surface with a positive slope. Here, φ_s is the angle made by the slope of the sea surface with the mean sea level, φ_1 and φ_2 are the angles made by the incident and reflected path with perpendicular to the mean sea level, R is the distance between Tx and Rx, and D_t and D_r are the depth of the Tx and Rx respectively with respect to the mean sea level. θ_c is AoA at the receiver. For negative slope, m, and in turn, φ_s becomes negative, and the overall algorithm remains the same. Let, X_{c_n} and Y_{c_n} represent the random variables that model the randomness in scatterer location and surface wave height, respectively, for n-th realization of the sea surface. Accordingly, X_c and Y_c represent the randomness in scatterer location and surface wave height for ensemble-averaged SoS surface.

A. Probability Density Function (PDF) of X_c

In order to model the symmetry inherent to each specular path, we propose X_{c_n} to be represented by a Gaussian Mixture Model (GMM) such that ([25], ch. 9)

$$f_{X_{c_n}}(X_{c_n}) = \frac{1}{N_s} \sum_{j=1}^{N_s} \mathcal{N}(X_{c_n} | \zeta_j, \eta), \tag{5}$$

where, N_s is the number of scatterer coordinates obtained for the n-th realization of sea surface. These scatterer locations are represented as ζ_j, $\{j = 1, 2, \ldots, N_s\}$ and have been computed using algorithm 1. η is the standard deviation (SD) of each component of the mixture model [26], [27]. We work with $\eta = 0.1$ m throughout the analysis. Fig. 4 gives a plot of the specular paths for the n-th realization of the sea surface. The value of the sea surface parameters obtained using MEA (as discussed in the Section II) are specified in the figure legend. At inset is the PDF of X_{c_n} using (5) obtained for a given geometry. The mean and variance of the X_{c_n} can be expressed as ([25], ch.9)

$$\mu_{X_{c_n}} = \sum_{j=1}^{N_s} v_j \zeta_j, \quad \sigma^2_{X_{c_n}} = \sum_{j=1}^{N_s} v_j (\eta^2 + \zeta_j^2) - \mu^2_{X_{c_n}}. \tag{6}$$

0018-9545 (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 December 06,2021 at 06:38:40 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/TVT.2021.3131504, IEEE Transactions on Vehicular Technology

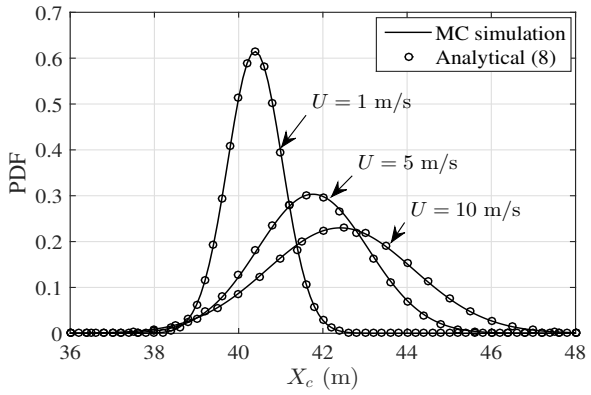


Fig. 5. PDF of X_c for a sea surface represented by SoS, obtained after averaging over $N = 40$ realizations, at $R = 100$ m, $D_t = 4$ m, $D_r = 6$ m, and $f = 2000$ Hz.

We assume all the paths to be equiprobable such that $v_j = 1/N_s$. Thus, total randomness of X_c for a sea surface represented by SoS can be expressed as

$$X_c = \frac{1}{N}\sum_{n=1}^{N} X_{c_n}. \tag{7}$$

This is equivalent to averaging the PDF of X_{c_n} over N realizations of the sea surface. The averaging leads to the convolution of PDF of $X_{c_n} = X_c|\phi'_n$ for all values of ϕ'_n. By central limit theorem, $X_c \sim \mathcal{N}(\mu_{X_c}, \sigma_{X_c})$ where, μ_{X_c} and σ_{X_c} are given by

$$\mu_{X_c} = \frac{1}{N}\sum_{n=1}^{N}\mu_{X_{c_n}}, \quad \sigma_{X_c} = \frac{1}{N}\sqrt{\sum_{n=1}^{N}\sigma^2_{X_{c_n}}}. \tag{8}$$

Fig. 5 gives a plot of the PDF of X_c for a sea surface represented by SoS at different values of wind speed. The plot is obtained after averaging over $N = 40$ realizations. The PDF is plotted using Monte Carlo (MC) simulation and analytical results using (8). We observe that as the wind speed increases, the correlation length of the sea surface increases. This leads to the enhanced spread of the scatterer locations on the surface. The increase in spread reflects in the PDF of X_c. We further observe that at $U = 1$ m/s, the mean of X_c lies close to the center of the geometry (scatterer location had the surface been flat). This is because at $U = 1$ m/s, the RMS wave height is very small (≈ 0.53 cm). For a Tx/Rx depth of few meters, the surface appears smooth. As the wind speed increases, the surface fluctuations increase due to which randomness in the scatterer position increases. Since we consider only forward scattering, the average location of the scatterer coordinates shifts to a higher value. Thus, the peak of the PDF shifts for increasing value of wind speeds.

B. PDF of Y_c

We observe that Y_{c_n} has induced randomness due to random behavior of X_{c_n} as well as ϕ'_n. PDF of Y_c can be obtained by summing up the randomness in Y_{c_n} for N realizations of ϕ'. Thus, we start with computing the PDF of $Y_{c_n} = Y_c|\phi'_n$ using (4). Let $\bar{X}_{c_n} = k_n X_{c_n} + \phi'_n$, then the PDF of $\bar{X}_{c_n}$ can be written as

$$f_{\bar{X}_{c_n}}(\bar{X}_{c_n}) = \frac{1}{N_s}\sum_{j=1}^{N_s}\mathcal{N}\left(\bar{X}_{c_n}|k_n\zeta_j + \phi'_n, k_n\eta\right). \tag{9}$$

Since $Y_{c_n} = a_n \sin\bar{X}_{c_n}$ is a many-to-one mapping, the roots of $\bar{X}_{c_n}$, given by $\bar{X}_{c_{np}}$ and $\bar{X}'_{c_{np}}$, can be expressed as

$$\left.\begin{aligned}\bar{X}_{c_{np}} &= (2p-2)\pi + \sin^{-1}\left(Y_{c_n}/a_n\right)\\ \bar{X}'_{c_{np}} &= (2p-1)\pi - \sin^{-1}\left(Y_{c_n}/a_n\right)\end{aligned}\right\} \text{ if } Y_{c_n} > 0, \tag{10}$$

$$\left.\begin{aligned}\bar{X}_{c_{np}} &= (2p-1)\pi + \sin^{-1}\left(Y_{c_n}/a_n\right)\\ \bar{X}'_{c_{np}} &= 2p\pi - \sin^{-1}\left(Y_{c_n}/a_n\right)\end{aligned}\right\} \text{ if } Y_{c_n} < 0, \tag{11}$$

where, $p = 1, 2, \ldots, N_p$ such that $N_p = \text{round}\left(Rk_n/2\pi\right)$[1] corresponds to full cycles of sine wave that would exist for a given range R and correlation length $2\pi/k_n$. PDF of Y_{c_n} can thus be expressed as

$$f_{Y_{c_n}}(Y_{c_n}) = \frac{1}{\sqrt{a_n^2 - Y_{c_n}^2}}\sum_{p=1}^{N_p}\left[f_{\bar{X}_{c_n}}(\bar{X}_{c_{np}}) + f_{\bar{X}_{c_n}}(\bar{X}'_{c_{np}})\right], \tag{12}$$

for $|Y_{c_n}| \leq |a_n|$. The PDF of Y_{c_n} obtained in (12) is 'U' shaped with maxima lying at $\pm a_n$. This implies that for a given realization of sea surface represented by a sinusoid, maximum scattering would occur either from the crest or trough. The PDF of Y_c can be obtained by convolving the PDF of Y_{c_n} for N realizations of ϕ' i.e.

$$f_{Y_c}(Y_c) = \bigotimes_{n=1}^{N} f_{Y_{c_n}}(Y_{c_n}), \quad -\sqrt{2}h \leq Y_c \leq \sqrt{2}h. \tag{13}$$

Here, $\bigotimes$ is the convolution operator. By central limit theorem, we conclude that Y_c is Gaussian distributed. Fig. 6 gives a plot of PDF of Y_c obtained using simulation (13) and the analytical expression given in [7], for different values of wind speed. For a range of $R = 100$ m, the mean of the surface wave height lies at the mean sea level. Thus, the peak of Fig. 6 does not shift with increasing wind speed. The exact match of the simulation result with the analytical result validates the analysis done in this Section. We also observe that the variance of Y_c increases with an increase of U. This can be attributed to an increase in the RMS wave height with increasing U. Thus, we can conclude that $Y_c \sim \mathcal{N}(\mu_{Y_c}, \sigma_{Y_c}) = \mathcal{N}(0, \sqrt{m_0/2\pi})$, where, m_0 is defined in Section II.

C. PDF of the AoA at the Rx

We now try to evaluate the PDF of AoA at the Rx for an ensemble-averaged SoS surface. From Fig. 1, AoA at the Rx (θ_c) for an averaged sea surface can be expressed as

$$\theta_c = \tan^{-1}\left(\frac{Y_c - y_r}{x_r - X_c}\right). \tag{14}$$

[1]round(Q): rounds each element of Q to the nearest integer.

0018-9545 (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 December 06,2021 at 06:38:40 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/TVT.2021.3131504, IEEE Transactions on Vehicular Technology

$$f_{\theta_c}(\theta_c) = \frac{R\sec^2\theta_c}{\sqrt{2\pi\left(\sigma_{X_c}^2\tan^2\theta_c + \sigma_{Y_c}^2\right)}} \exp\left[-\frac{\left(R\tan\theta_c + y_r - \mu_{X_c}\tan\theta_c - \mu_{Y_c}\right)^2}{2\left(\sigma_{X_c}^2\tan^2\theta_c + \sigma_{Y_c}^2\right)}\right], \quad 0 \le \theta_c \le \frac{\pi}{2} \tag{18}$$

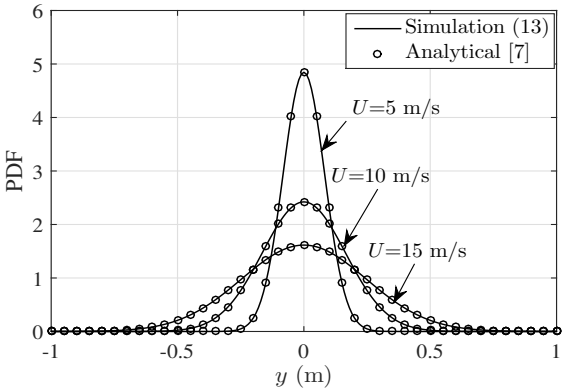


Fig. 6. PDF of Y_c obtained after averaging with $N = 40$ realizations of ϕ' at different wind speeds.

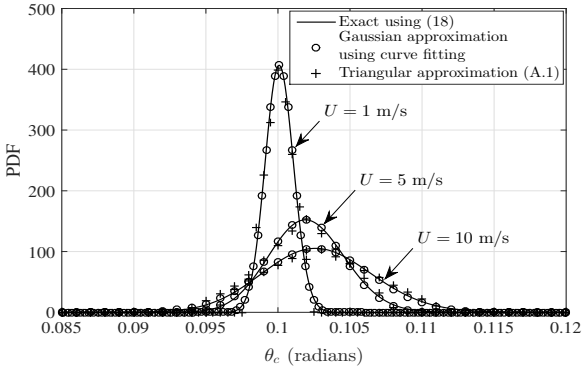


Fig. 7. PDF of the AoA at the Rx for coherent scattering for different values of wind speeds at $R = 100$ m, $D_t = 4$ m, $D_r = 6$ m, and $f = 2000$ Hz.

Based on the above discussions, we know that X_c and Y_c are Gaussian distributed with mean and variances determined by geometry and sea surface conditions. In terms of probability, we can express the CDF of θ_c as

$$\begin{aligned}\Pr(\theta_c \le \theta) &= \Pr\left(\tan^{-1}\left(\frac{Y_c - y_r}{x_r - X_c}\right) \le \theta\right) \\ &= \Pr(Y_c + X_c\tan\theta \le x_r\tan\theta + y_r). \end{aligned} \tag{15}$$

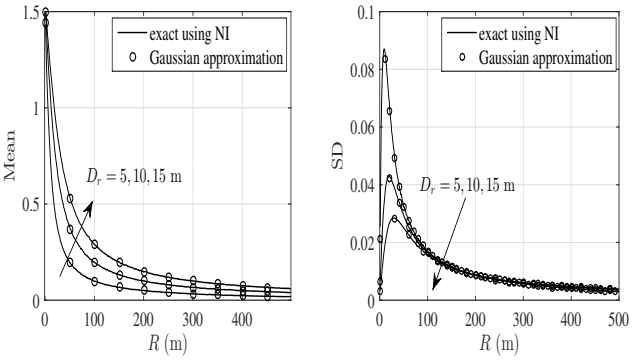
If we substitute $Z = Y_c + X_c\tan\theta$, CDF can be expressed as $F_{\theta_c}(\theta) = F_Z(x_r\tan\theta + y_r)$. Accordingly, the PDF of AoA is given as

$$f_{\theta_c}(\theta) = x_r\sec^2\theta f_Z(x_r\tan\theta + y_r). \tag{16}$$

Since X_c and Y_c are independent random variables, we can write

$$Z \sim \mathcal{N}\left(\mu_{X_c}\tan\theta + \mu_{Y_c}, \sqrt{\sigma_{X_c}^2\tan^2\theta + \sigma_{Y_c}^2}\right). \tag{17}$$

The PDF of θ_c can thus be expressed as (18). Fig. 7 gives the plot of the PDF of θ_c for a given geometry for different values of wind speed. We observe that at $U = 1$ m/s, when the surface is almost flat for the given depth of Tx or Rx, PDF of θ_c is symmetrical around the mean $\mu_c = \tan^{-1}\left(\frac{D_t+D_r}{R}\right)$. As U increases, the mean of θ_c shifts towards the right, and the PDF becomes slightly positively skewed. However, the skewness is so small (0.017 at $U = 10$ m/s) that a normal distribution can approximate the PDF. A reasonable match of the Gaussian approximation with the exact expression in Fig. 7 for different wind speeds validates the approximation. Thus, we conclude that $\theta_c \sim \mathcal{N}(\mu_c, \sigma_c)$, where, μ_c and σ_c can be computed using curve fitting. We also observe that due to the small variance, the triangular approximation (Appendix A) also provides a close match to the PDF of θ_c. This approximation is used to enhance the mathematical tractability of PDF of θ_c in Section VI. The parameters for the triangular approximation are computed using (A.2).



(a) Mean (b) Standard Deviation

Fig. 8. Mean and SD of the AoA at the Rx due to coherent scattering, as a function of R and D_r, at $U = 20$ m/s.

Fig. 8 demonstrates the impact of geometry on the parameters of θ_c. It gives a plot of the mean and SD of θ_c for increasing values of R at different values of D_r. Deriving a closed-form expression for mean and SD of θ_c is mathematically intractable. Thus, the exact value of parameters is obtained using numerical integration (NI). For ease of implementation and understanding, we keep $D_t = D_r$. From Fig. 8(a), we observe that the mean decreases with increasing R and increases with increasing D_r. In Fig. 8(b), for small values of R, when the Tx and Rx are very close to each other, the small region of the sea surface between Tx and Rx produces limited specular paths leading to a small value of SD. As R increases, the region between Tx and Rx increases, leading to increased randomness, and hence SD also increases. For large range, the correlation length of the sea surface relative

0018-9545 (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 December 06,2021 at 06:38:40 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/TVT.2021.3131504, IEEE Transactions on Vehicular Technology

to the position of Tx and Rx becomes very small that the surface appears smooth, leading to reduced SD. Further, as D_r increases, the randomness of the sea surface relative to Tx/Rx depth decreases, and hence SD decreases. The error in approximation of the PDF of θ_c would increase with increasing value of wind speed. Thus, the plot is obtained at $U = 20$ m/s. The precise match of the parameters retrieved using Gaussian approximation with the parameters obtained using NI validates the approximation for different geometries and large wind speeds.

IV. AoA distribution at the receiver for incoherent scattering

When the sea surface roughness increases, the scattering pattern at the surface changes. At relatively higher frequencies and wind speeds, every point on the surface acts as a potential scatterer. This phenomenon is called incoherent scattering, where the phase of the scattered signal is highly random. Some part of the scattered component gets canceled due to random phase. However, a large part remains that contribute to the scattered field [8], [11], [28]. In this section, we evaluate the randomness of potential scatterers on the sea surface represented by X_{ic}, randomness in surface wave height represented by Y_{ic}, and the distribution of AoA at the receiver (θ_{ic}), for incoherent scattering.

A. PDF of X_{ic}

For a given geometry and acoustic signal frequency, at large values of wind speed, the RMS wave height of the sea surface increases, and it becomes rough. An acoustically rough surface would result in scattering from every point along the surface [23]. This phenomenon can be interpreted as points of interaction or scatterers, which are distributed uniformly along the surface [9], [22]. Thus, the distribution of X_{ic} can be expressed as

$$f_{X_{ic}}(X_{ic}) = \frac{1}{R}, \quad 0 \le X_{ic} \le R. \tag{19}$$

B. PDF of Y_{ic}

Using the same procedure as in previous section, we start with observing that $X_{ic} \sim \mathcal{U}(0, R) \implies \bar{X}_{ic} = k_n X_{ic} + \phi'_n \sim \mathcal{U}(\phi'_n, k_n R)$. The mapping from $\bar{X}_{ic}$ to Y_{ic_n} is many-to-one and thus using (12), PDF of Y_{ic_n} for a given realization of sea surface can be expressed as

$$f(Y_{ic_n}) = \frac{N_p}{R k_n \sqrt{a_n^2 - Y_{ic_n}^2}}, \quad |Y_{ic_n}| \le |a_n|. \tag{20}$$

The overall PDF of Y_{ic} can be evaluated using (13). Because of central limit, Y_{ic} for incoherent scattering is again Gaussian distributed with parameters given as $Y_{ic} \sim \mathcal{N}(\mu_{Y_{ic}}, \sigma_{Y_{ic}}) = \mathcal{N}(0, \sqrt{m_0/2\pi})$.

Remark 1. *The analysis for distribution of surface wave height for coherent and incoherent scattering indicates that, irrespective of the kind of scattering that occurs in an UW channel, sea surface represented by SoS implies Gaussian surface roughness. Most of the experiments conducted at different geographical locations suggest Gaussian distribution of sea surface wave height [4], ([29], ch. 3). Hence, this work can be applied to a wide range of water bodies.*

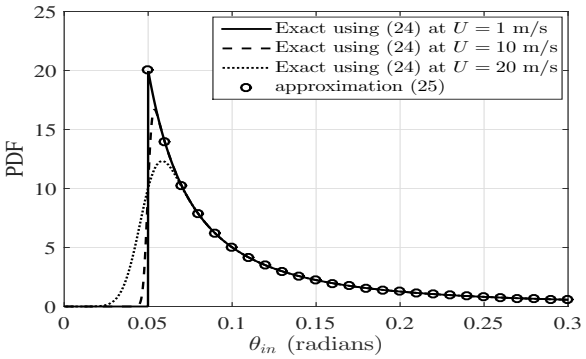


Fig. 9. PDF of AoA at the Rx due to incoherent scattering at $R = 100$ m, $D_t = D_r = 5$ m at different values of wind speeds.

C. PDF of the AoA at the Rx

If the AoA at the receiver for incoherent scattering from the sea surface is represented as θ_{ic}, it follows from (14) that

$$\begin{aligned}\Pr(\theta_{ic} \le \theta) &= \Pr\left(\tan^{-1}\left(\frac{Y_{ic} - y_r}{x_r - X_{ic}}\right) \le \theta\right) \\ &= \Pr\left(Y_{ic}\cot\theta + X_{ic} \le x_r + y_r \cot\theta\right). \end{aligned} \tag{21}$$

Making the substitution $Z' = Y_{ic}\cot\theta + X_{ic}$, one obtains

$$\begin{aligned} F_{\theta_{ic}}(\theta) &= F_{Z'}\left(x_r + y_r\cot\theta\right) \\ f_{\theta_{ic}}(\theta_{ic}) &= -y_r \csc^2\theta_{ic} f_{Z'}\left(x_r + y_r\cot\theta_{ic}\right) \end{aligned} \tag{22}$$

From the above analysis, we conclude that $Y'_{ic} = Y_{ic}\cot\theta \sim \mathcal{N}(0, \cot\theta\sqrt{m_0/2\pi})$. We can write

$$\begin{aligned} f_{Z'}(Z') &= \int_0^R f_{Y'_{ic}}(Z' - u) f_x(u) du \\ &= \frac{1}{\sqrt{2\pi\sigma^2_{Y'_{ic}}}R}\int_0^R \exp\left[-\frac{(Z'-u)^2}{2\sigma^2_{Y'_{ic}}}\right] du \\ &= \frac{1}{2R}\left[\text{Erf}\left[\frac{-Z'+R}{\sqrt{2}\sigma_{Y'_{ic}}}\right] + \text{Erf}\left[\frac{Z'}{\sqrt{2}\sigma_{Y'_{ic}}}\right]\right]. \end{aligned} \tag{23}$$

where, $\text{erf}(\nu) = \frac{2}{\sqrt{\pi}}\int_0^\nu e^{-t^2} dt$. Using (22) and (23), the PDF of θ_{ic} can be expressed as (24). Since, $\sigma_{Y_{ic}} \ll 1$ even for $U = 15$ m/s, error function $\text{Erf}\left[\frac{R\tan\theta_{ic} - D_r}{\sqrt{2}\sigma_{Y_{ic}}}\right]$ appears almost like a signum function with discontinuity at $\theta_{ic} = \tan^{-1}\left(\frac{D_r}{R}\right)$. Thus, PDF in (24) can be reduced to

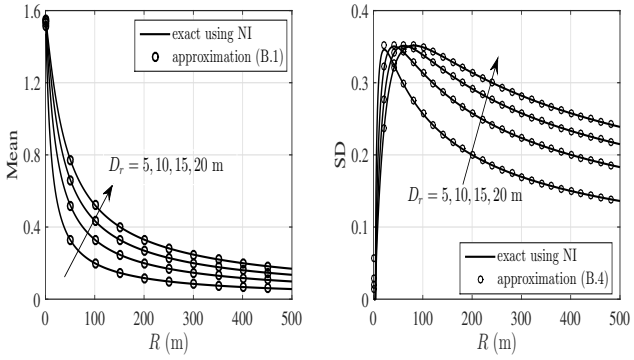
$$f^{\text{approx}}_{\theta_{ic}}(\theta_{ic}) = \frac{D_r}{R\sin^2\theta_{ic}}, \quad \tan^{-1}\left(\frac{D_r}{R}\right) \le \theta_{ic} \le \frac{\pi}{2}, \tag{25}$$

for low wind speeds. This is the exact expression as derived in [22] for a flat sea surface. The expression in (24) takes care of discontinuity that exists in [22] due to minimum sea surface fluctuations. When the fluctuations are zero, the PDF

0018-9545 (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 December 06,2021 at 06:38:40 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/TVT.2021.3131504, IEEE Transactions on Vehicular Technology

$$f_{\theta_{ic}}(\theta_{ic}) = \frac{D_r}{2R\sin^2\theta_{ic}}\left[\text{Erf}\left[\frac{D_r}{\sqrt{2}\sigma_{Y_{ic}}}\right] + \text{Erf}\left[\frac{R\tan\theta_{ic} - D_r}{\sqrt{2}\sigma_{Y_{ic}}}\right]\right], \ \tan^{-1}\left(\frac{D_r - \sqrt{2}h}{R}\right) \leq \theta_{ic} \leq \frac{\pi}{2} \tag{24}$$



(a) Mean (b) Standard Deviation

Fig. 10. Mean and SD of the AoA at the Rx for incoherent scattering as a function of R and D_r at $U = 20$ m/s.

in [22] is valid. However, as fluctuations increase, the value of $\sigma_{Y_{ic}}$ increase linearly making Erf$\left[\frac{R\tan\theta_{ic}-D_r}{\sqrt{2}\sigma_{Y_{ic}}}\right]$ deviate from signum behavior. This makes the PDF of θ_{ic} continuous. Fig. 9 gives a plot of PDF of θ_{ic} using (24) and (25) for different values of wind speed and small value of D_r. We observe that the PDF of θ_{ic} gradually becomes continuous for increasing values of U.

Fig. 10 gives a plot of mean and SD of θ_{ic} for increasing values of R and D_r. The exact value is obtained using NI, and approximate parameters are derived in Appendix B. From the derived expressions, we infer that maximum error in approximation would appear for a large value of U and for small value of R and D_r. To get an idea of the bound on error, we have computed the parameters at $U = 20$ m/s. We observe that the mean of the θ_{ic} obtained using approximation matches reasonably well with the NI results. A small approximation error is observed for a small value of R and D_r in the plot of SD, which gradually reduces and becomes almost negligible for larger ranges. For smaller values of U, the parameters would overlap, as is also evident from the expressions.

V. PDF OF OVERALL AOA

The overall AoA at any given geometry and wind speed can be expressed as a function of both coherent and incoherent AoA [8], [30]. The transformation from coherent to incoherent is gradual and is a function of acoustic signal frequency, wind speed, and geometry. It was proposed in [30], and later followed by several authors such as [31], that the measured field in each experiment is the sum of coherent average and the incoherent field. Thus, the overall AoA (θ) for any given surface roughness can be expressed as

$$\theta = \rho_c\theta_c + \rho_{ic}\theta_{ic} = \hat{\theta}_c + \hat{\theta}_{ic}, \tag{26}$$

where, ρ_c is the coherent intensity reflection coefficient (RC) of the acoustic signal, and $\rho_{ic} = 1 - \rho_c$. This ensures that the scattering is completely coherent when the surface is smooth and a combination of coherent and incoherent otherwise [8]. Various empirical models for RC are available in literature [5]. For a sea surface with Gaussian distribution of surface wave height, Kirchhoff approximation (KA) to coherent RC is given by [5], [8]

$$\rho_c = \exp\left(-\chi^2/2\right), \tag{27}$$

where, χ is the Rayleigh parameter. Using observations made in the previous sections, we can write $\hat{\theta}_c \sim \mathcal{N}(\rho_c\mu_c, \rho_c\sigma_c) \sim \mathcal{N}(\hat{\mu}_c, \hat{\sigma}_c)$, in the range $0 \leq \hat{\theta}_c \leq \rho_c\pi/2$. Also since, $\hat{\theta}_{ic} = \rho_{ic}\theta_{ic}$ is a one-to-one function for a given wind speed and frequency of the acoustic signal, the PDF of $\hat{\theta}_{ic}$ is given by $f_{\hat{\theta}_{ic}}(\hat{\theta}_{ic}) = \frac{1}{|\rho_{ic}|}f_{\theta_{ic}}\left(\frac{\hat{\theta}_{ic}}{\rho_{ic}}\right)$ ([32], ch. 5). Since, the PDF of θ_c and θ_{ic} are independent of each other, the overall PDF of θ is convolution of $\hat{\theta}_c$ and $\hat{\theta}_{ic}$ such that

$$f_\theta(\theta) = f_{\hat{\theta}_c}(\theta) \bigotimes f_{\hat{\theta}_{ic}}(\theta) = \int f_{\hat{\theta}_c}(\Phi) f_{\hat{\theta}_{ic}}(\theta - \Phi) d\Phi. \tag{28}$$

PDF of θ thus obtained using exact and approximate expressions for θ_c and θ_{ic} is given by (29) and (30). Here, $\Phi_1 = \theta - \frac{\rho_{ic}\pi}{2}$, $\Phi_2 = \min\left(\frac{\rho_c\pi}{2}, \theta - \rho_{ic}\tan^{-1}\frac{D_r-\sqrt{2}h}{R}\right)$, and $\Phi_3 = \min\left(\frac{\rho_c\pi}{2}, \theta - \rho_{ic}\tan^{-1}\frac{D_r}{R}\right)$. Obtaining a closed-form expression for the PDF of θ using (29) and (30) appears intractable. Thus we use triangular approximation to the Gaussian distribution as defined in Appendix A. The closed-form expression for the PDF of θ thus obtained is given by (33). The parameters of the overall AoA can be computed as

$$\begin{cases} \mu_\theta = E[\theta] = \rho_c\mu_c + (1-\rho_c)\mu_{in} \\ \sigma_\theta = E[(\theta - \mu_\theta)^2] = \sqrt{\rho_c^2\sigma_c^2 + (1-\rho_c)^2\sigma_{in}^2}. \end{cases} \tag{31}$$

VI. RESULTS AND DISCUSSIONS

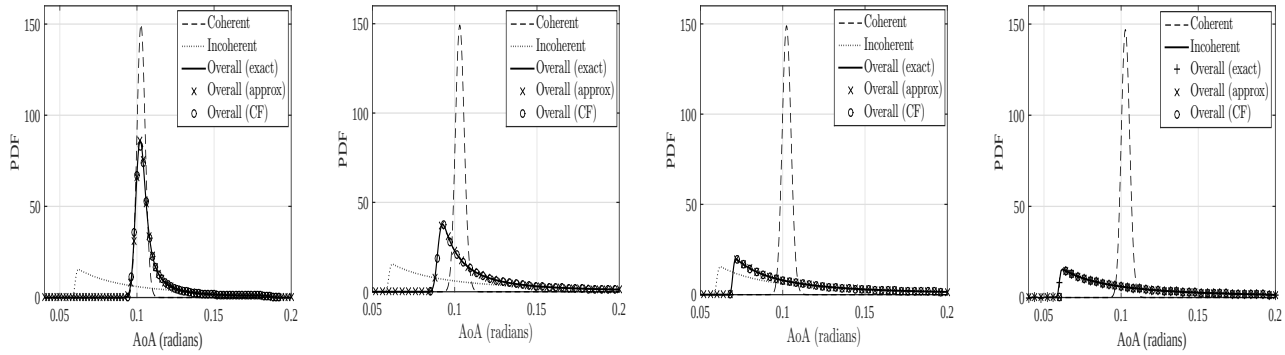
In this section, we study the impact of acoustic signal frequency, the UW channel geometry, and the wind speed on the PDF of overall AoA and its parameters. To study the impact of sea surface fluctuations effectively, we choose small depths of Tx and Rx (few meters). Note that for a given time instant, v_s determines the wavelength of the surface wave. An increase in v_s would lead to enhanced surface wavelength and reduced fluctuations. The surface wavelength would further increase with wind speed for a given value of v_s. This is because the power spectral density of a JONSWAP spectrum model shifts to a lower frequency range for higher wind speeds. Since the focus of this work is to evaluate the impact of sea surface fluctuations on the AoA, we carry out the analysis for a fixed value of $v_s = 1$ m/s at $t = 1$ sec. We consider $F = 100$ km throughout the analysis. All the angles are measured in radians. The channel is assumed to be isovelocity with $c = 1500$ m/s.

0018-9545 (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 December 06,2021 at 06:38:40 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/TVT.2021.3131504, IEEE Transactions on Vehicular Technology

$$f_\theta(\theta) = \frac{D_r}{2\rho_{ic}\rho_c}\int_{\Phi_1}^{\Phi_2} \frac{\sec^2\left(\frac{\Phi}{\rho_c}\right)}{\sqrt{2\pi\left(\sigma_{X_c}^2\tan^2\left(\frac{\Phi}{\rho_c}\right)+\sigma_{Y_c}^2\right)}}\exp\left[-\frac{\left((R-\mu_{X_c})\tan\left(\frac{\Phi}{\rho_c}\right)+y_r-\mu_{Y_c}\right)^2}{2\left(\sigma_{X_c}^2\tan^2\left(\frac{\Phi}{\rho_c}\right)+\sigma_{Y_c}^2\right)}\right]$$
$$\times\frac{1}{\sin^2\left(\frac{\theta-\Phi}{\rho_{ic}}\right)}\left[\text{Erf}\left[\frac{D_r}{\sqrt{2}\sigma_{Y_c}}\right]+\text{Erf}\left[\frac{R\tan\left(\frac{\theta-\Phi}{\rho_{ic}}\right)-D_r}{\sqrt{2}\sigma_{Y_c}}\right]\right]d\Phi,\quad \tan^{-1}\left(\frac{D_r-\sqrt{2}h}{R}\right)\le\theta\le\frac{\pi}{2} \tag{29}$$

$$f_\theta^{\text{approx}}(\theta) = \frac{D_r}{R|\rho_{ic}|\sqrt{2\pi\rho_c^2\sigma_c^2}}\int_{\Phi_1}^{\Phi_3}\exp\left[-\frac{(\Phi-\rho_c\mu_c)^2}{2\rho_c^2\sigma_c^2}\right]\frac{1}{\sin^2\left(\frac{\theta-\Phi}{\rho_{ic}}\right)}d\Phi,\quad \tan^{-1}\left(\frac{D_r}{R}\right)\le\theta\le\frac{\pi}{2} \tag{30}$$



(a) $f = 2000$ Hz (b) $f = 5000$ Hz (c) $f = 10$ kHz (d) $f = 20$ kHz

Fig. 11. PDF of AoA at the Rx due to coherent, incoherent, and overall scattering. The result for overall scattering is obtained using the exact (29), approximate (30), and closed-form expressions (33) at $R = 100$ m, $D_t = 4$ m, $D_r = 6$ m, and $U = 5$ m/s.

A. Effect of the frequency of the acoustic signal

As the demand for data rate and capacity increases, there is growing attention on the high-frequency analysis of the UW acoustic channel. Thus, we study the impact of using different acoustic frequencies on the AoA at the Rx. Figs. 11(a)-11(d) give the plot of PDF of AoA due to coherent, incoherent, and overall scattering at different value of acoustic frequencies for a given geometry and wind speed. We observe that f does not have any impact on the parameters of θ_c and θ_{ic}. However, for small values of f, the overall AoA PDF is mostly coherent. As the acoustic frequency increases, the PDF of overall AoA gradually shifts from coherent to incoherent. At $f = 20$ kHz, the scattering from the sea surface at a nominal wind speed of $U = 5$ m/s is mostly incoherent. Thus, we conclude that at higher frequencies, most naturally occurring sea surfaces result in incoherent scattering of the acoustic signal. This observation is in agreement with [6], [11], [23], where the frequency of operation is large and scattering is dominated by incoherent part for any geometry or wind speed.

B. Effect of geometry of an UW channel

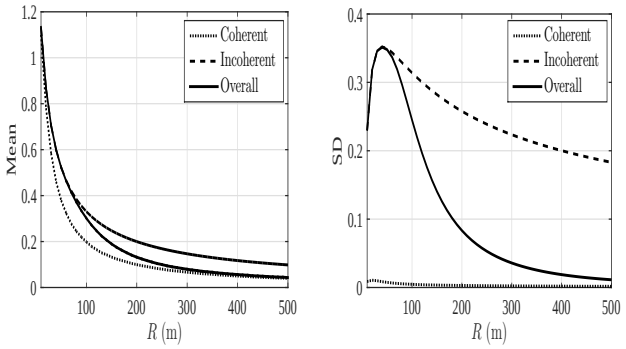
Figs. 12 and 13 plot the mean and SD of the coherent, incoherent, and overall AoA as a function of geometry. For these figures, we assume $D_r = D_t$ for ease of implementation and understanding. From Fig. 12, we observe that for a given wind speed, as R increases, both mean and SD of all the AoA decrease. The parameters of the overall AoA, however, gradually shift from incoherent to coherent for increasing R. This can be attributed to the reduced grazing angle for large ranges, that make the Rayleigh parameter small. This shift from incoherent to coherent scattering is a function of wind speed and acoustic signal frequency. Thus, we conclude that as R increases, the scattering gradually becomes coherent. On the other hand, in Fig. 13, as D_r or D_t increase, both mean and SD increase. The parameters of the overall AoA shift from coherent to incoherent and scattering eventually become incoherent, as the grazing angle increases. The parameters for coherent scattering are obtained using Gaussian approximation, and the parameters for incoherent scattering are obtained using (B.1) and (B.4), as discussed in Sections III and IV respectively. The parameters for the overall scattering are obtained using (31).

C. Effect of wind speed

Figs. 14(a)-14(d) plot the PDF of AoA at the Rx for coherent, incoherent, and overall scattering at different values of wind speed. We observe that at $U = 1$ m/s, the PDF of the overall AoA almost overlap with the PDF of coherent AoA. This is because at $U = 1$ m/s, sea surface roughness is very small (as discussed in Section III A). For a Tx/Rx

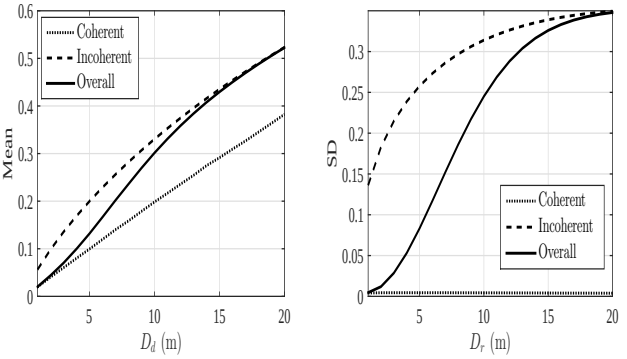
0018-9545 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 06,2021 at 06:38:40 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/TVT.2021.3131504, IEEE Transactions on Vehicular Technology



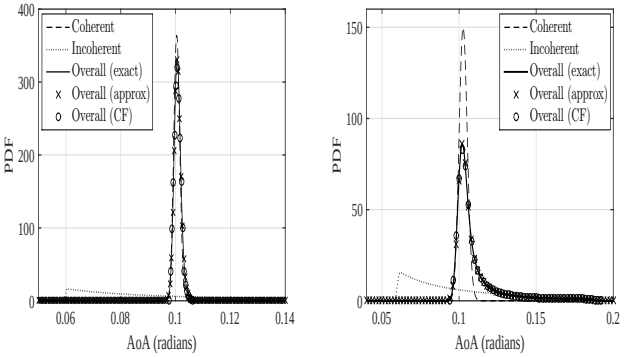
(a) Mean (b) Standard Deviation

Fig. 12. Mean and SD of the coherent, incoherent, and overall AoA at $D_t = D_r = 10$ m, $U = 10$ m/s, and $f = 2000$ Hz.

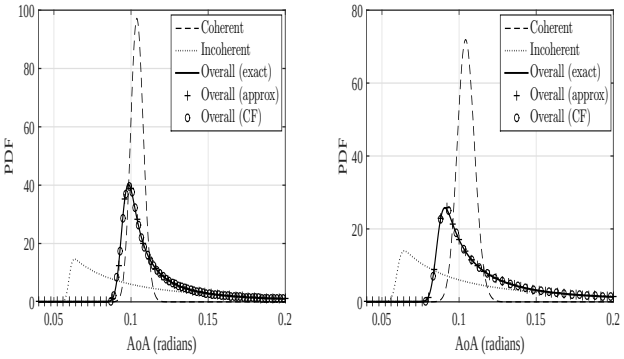


(a) Mean (b) Standard Deviation

Fig. 13. Mean and SD of the coherent, incoherent, and overall AoA at $R = 100$ m, $U = 10$ m/s, and $f = 2000$ Hz.



(a) $U = 1$ m/s (b) $U = 5$ m/s



(c) $U = 10$ m/s (d) $U = 15$ m/s

Fig. 14. PDF of AoA at the Rx due to coherent, incoherent, and overall scattering. The result for overall scattering is obtained using the exact (29), approximate (30), and closed-form expressions (33) at $R = 100$ m, $D_t = 4$ m, $D_r = 6$ m, and $f_c = 2000$ Hz.

depth of few meters, the surface appears almost flat. Thus, scattering is entirely coherent. As U increases, the surface roughness increases due to increased RMS wave height. This leads to the PDF of overall AoA gradually shifting towards the incoherent component. The magnitude of the shift is a function of geometry and acoustic signal frequency, as discussed above.

Fig. 15 gives the plot of the SD of coherent, incoherent, and the overall AoA as a function of wind speed. We observe that SD of the coherent AoA increase with U. This increase, however, is a function of geometry, as depicted in Fig. 8. Parameters of incoherent AoA are nearly independent of the wind speed. The SD of the overall AoA gradually shifts from coherent to incoherent with increasing U, which is in agreement with Fig. 14. The magnitude of the shift is a function of geometry, acoustic signal frequency, and the value of RC preferred. We have given the result for KA to RC given in [5] and the small slope approximation (SSA) model for the RC at small grazing angles given in [33]. It was observed in [5] that reflection loss obtained using the SSA model is closer to the actual loss as compared to loss due to the KA model at lower grazing angles. Different values of RC may be chosen for different geometries and sea surface conditions and is a subject of open research.

Remark 2. *The exact match of the overall AoA PDF obtained using the exact expressions* (29)*, approximation* (30)*, and the closed-form expression* (33) *at different frequencies and wind speeds in Figs. 11 and 14, validate the approximation and closed-form expression.*

Further, we would like to mention that the AoA modeling proposed in this work provides an effective mechanism to capture the time-variation of the channel. As the Tx/Rx moves (intentional or unintentional), the AoA changes, the effect of which reflects on the statistical characteristics of the channel. This can be verified by the plot of Doppler spectrum in Fig. 16 for a Rx moving at a speed of $v_r = 1$ m/s towards the Tx. It is shown in [19] that for an UW channel, approximate Doppler power spectrum can be calculated directly from the PDF of AoA as $S_{hh}(f) = \dfrac{f_\theta(\cos^{-1}(f/f_m)) + f_\theta(-\cos^{-1}(f/f_m))}{\sqrt{f_m^2 - f^2}}$ for $|f| \leq f_m$. Here, $f_m = \dfrac{v_r}{c} f_c$ is the Doppler shift produced by relative Tx/Rx motion. The plot is obtained at different values of wind speed above the sea surface. We observe that Doppler spread increases with increasing wind speed. This is intuitive as increased value of U implies enhanced sea surface randomness which increases the spread. Note that the amount of Doppler spread produced for a given relative speed of the Tx/Rx and acoustic frequency would thus be a function of

0018-9545 (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 December 06,2021 at 06:38:40 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/TVT.2021.3131504, IEEE Transactions on Vehicular Technology

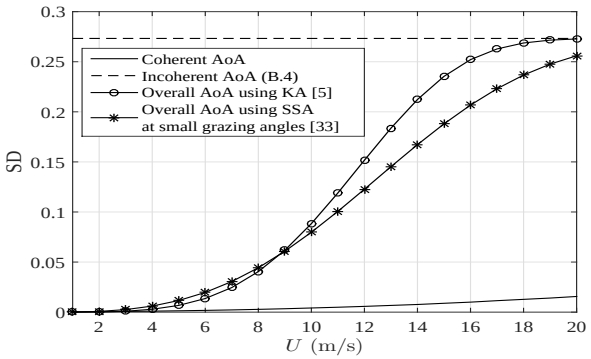


Fig. 15. A comparison of the SD as a function of wind speed, for coherent, incoherent and overall AoA at $R = 100$ m, $D_t = 4$ m, $D_r = 6$ m, and $f = 2000$ Hz.

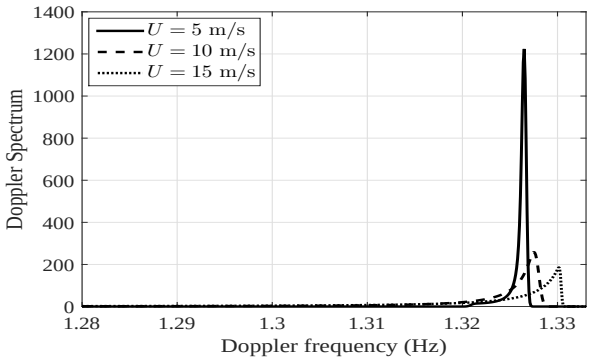


Fig. 16. Doppler spectrum at the Rx for $v_r = 1$ m/s, $R = 100$ m, $D_r = 4$ m, $D_t = 6$ m, and $f = 2000$ Hz, at different values of wind speed.

wind speed and geometry in an UW channel.

VII. Conclusion

In this work, we have proposed a model for average AoA at the receiver due to scattering from a random sea surface. The sea surface is represented using the sum of sinusoids (which ensures Gaussian surface roughness) and modeled using the JONSWAP sea wave spectrum. Since most natural sea surfaces have been observed to have Gaussian surface roughness, we conclude that AoA modeling is general and widely applicable. To the author's best knowledge, this is the first attempt to compute the average AoA at the receiver for a fluctuating sea surface. The derived close-form expressions are comprehensive and a function of wind speed above the sea surface, acoustic signal frequency, channel geometry, and fetch limit range. The analysis performed in this work is based on acoustic surface roughness and can be applied to bottom scattering directly by characterizing the bottom profile. It can further be extended for multiple bounce paths. The AoA PDF so derived can be used to evaluate the Doppler spread due to vehicular motion in an UW channel. It can further be used for source localization or optimizing the angular positioning of a directional receiver in a communication network.

Appendix A
Closed-Form expression for the PDF of θ using Triangular approximation

Symmetric Triangular distribution is defined as ([34], ch. 40)

$$f_T(\psi) = \begin{cases} \frac{4(\psi-\psi_1)}{(\psi_3-\psi_1)^2} & \text{for} \quad \psi_1 \le \psi \le \psi_2 \\ \frac{4(\psi_3-\psi)}{(\psi_3-\psi_1)^2} & \text{for} \quad \psi_2 < \psi \le \psi_3 \\ 0 \quad \text{otherwise} \end{cases} \tag{A.1}$$

where, $\psi_2 = (\psi_1 + \psi_3)/2$ is the mean of the distribution. If a random variable $\hat{\theta}_c \sim \mathcal{N}(\hat{\mu}_c, \hat{\sigma}_c)$, it can be approximated with Triangular distribution PDF given in (A.1) using moment matching technique such that

$$\psi_1 = \hat{\mu}_c - \sqrt{6}\hat{\sigma}_c, \quad \psi_3 = \hat{\mu}_c + \sqrt{6}\hat{\sigma}_c \tag{A.2}$$

Using (30) and (A.1), PDF of θ can be expressed as (32) where, $L_1 = \max\left(\psi_1, \theta - \frac{\rho_{ic}\pi}{2}\right)$, $L_2 = \min\left(\psi_2, \theta - \rho_{ic}\tan^{-1}\frac{D_r}{R}\right)$, $L_3 = \max\left(\psi_2, \theta - \frac{\rho_{ic}\pi}{2}\right)$ and $L_4 = \min\left(\psi_3, \theta - \rho_{ic}\tan^{-1}\frac{D_r}{R}\right)$. After solving the integration, PDF of θ can be expressed as (33), where A_1 and B_1 are defined in (34).

Appendix B
Closed-form expressions for the Mean and Variance of θ_{ic}

1) Mean of θ_{ic}: The mean of the approximate PDF of θ_{ic} can be expressed as

$$\begin{aligned} \mu_{ic} = E[\theta_{ic}] &= \int_{\tan^{-1}\left(\frac{D_r}{R}\right)}^{\pi/2} \theta_{ic} f_{\theta_{ic}}(\theta_{ic}) d\theta_{ic} \\ &= \frac{D_r}{R}\int_{\tan^{-1}\left(\frac{D_r}{R}\right)}^{\pi/2} \theta_{ic}\csc\theta_{ic}^2 d\theta_{ic} \\ &= \tan^{-1}\left(\frac{D_r}{R}\right) - \frac{D_r}{R}\ln\left[\frac{D_r}{\sqrt{R^2+D_r^2}}\right] \end{aligned} \tag{B.1}$$

2) Variance of θ_{ic}: The variance of the approximate PDF of θ_{ic} can be expressed as

$$\sigma_{ic}^2 = E[(\theta_{ic} - \mu_{ic})^2] = E[\theta_{ic}^2] - \mu_{ic}^2$$

Here, the mean square value $E[\theta_{ic}^2]$ can be computed as

$$E[\theta_{ic}^2] = \frac{D_r}{R}\int_{\tan^{-1}\left(\frac{D_r}{R}\right)}^{\pi/2} \theta_{ic}^2\csc\theta_{ic}^2 d\theta_{ic} \tag{B.2}$$

Solving using integration by parts, we get (B.3). Here $\xi_{i_2}(x)$ is the polylogarithmic function [35]. The variance thus obtained is given by (B.4).

References

[1] M. R. Mousavi, M. Karimi, and A. Jamshidi, "Probability distribution of acoustic scattering from slightly rough sea surface," *Ocean Eng.*, vol. 112, pp. 134–144, 2016.
[2] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 84–89, 2009.
[3] L. Lachman, "An open programming architecture for modeling ocean waves," in *IMAGE Conference, Scottsdale, Arizona*, 2007, pp. 1–13.

0018-9545 (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 December 06,2021 at 06:38:40 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/TVT.2021.3131504, IEEE Transactions on Vehicular Technology

$$f_\theta(\theta) = \frac{4D_r}{R\rho_{ic}(\psi_3-\psi_1)^2}\left[\int_{\psi_1}^{\psi_2}\frac{(\Phi-\psi_1)}{\sin^2\left(\frac{\theta-\Phi}{\rho_{ic}}\right)}d\Phi + \int_{\psi_2}^{\psi_3}\frac{(\psi_3-\Phi)}{\sin^2\left(\frac{\theta-\Phi}{\rho_{ic}}\right)}d\Phi\right] = \frac{4D_r}{R\rho_{ic}(\psi_3-\psi_1)^2}\left[\int_{L_1}^{L_2}\frac{(\Phi-\psi_1)}{\sin^2\left(\frac{\theta-\Phi}{\rho_{ic}}\right)}d\Phi + \int_{L_3}^{L_4}\frac{(\psi_3-\Phi)}{\sin^2\left(\frac{\theta-\Phi}{\rho_{ic}}\right)}d\Phi\right] \quad (32)$$

$$f_\theta(\theta) = \frac{4}{(\psi_3-\psi_1)^2}\frac{D_r}{R}\times\begin{cases} A_1, & \text{for } \psi_1+\rho_{ic}\tan^{-1}\frac{D_d}{R} \le \theta \le \psi_2+\rho_{ic}\tan^{-1}\frac{D_d}{R} \\ A_1+B_1, & \text{for } \psi_2+\rho_{ic}\tan^{-1}\frac{D_r}{R} \le \theta \le \psi_2+\rho_{ic}\frac{\pi}{2} \\ B_1, & \text{for } \psi_2+\rho_{ic}\frac{\pi}{2} \le \theta \le \psi_3+\rho_{ic}\frac{\pi}{2} \end{cases} \quad (33)$$

$$\left.\begin{aligned} A_1 &= (-\psi_1+L_1)\cot\left(\tfrac{L_1-\theta}{\rho_{ic}}\right)+(\psi_1-L_2)\cot\left(\tfrac{L_2-\theta}{\rho_{ic}}\right)+\rho_{ic}\left\{-\log\left(-\sin\left(\tfrac{L_1-\theta}{\rho_{ic}}\right)\right)+\log\left(-\sin\left(\tfrac{L_2-\theta}{\rho_{ic}}\right)\right)\right\} \\ B_1 &= (\psi_3-L_3)\cot\left(\tfrac{L_3-\theta}{\rho_{ic}}\right)+(-\psi_3+L_4)\cot\left(\tfrac{L_4-\theta}{\rho_{ic}}\right)+\rho_{ic}\left\{\log\left(-\sin\left(\tfrac{L_3-\theta}{\rho_{ic}}\right)\right)+\log\left(-\sin\left(\tfrac{L_4-\theta}{\rho_{ic}}\right)\right)\right\} \end{aligned}\right\} \quad (34)$$

$$E\left[\theta_{ic}^2\right] = \frac{D_r}{R}\Big[\frac{-i\pi^2}{6}+\pi\log 2-\frac{1}{D_r}\Big\{\tan^{-1}\frac{D_r}{R}\Big((-R-iD_r)\tan^{-1}\frac{D_r}{R} + 2D_r\log\left(1-e^{2i\tan^{-1}\frac{D_r}{R}}\right)\Big)-iD_r\xi_{i_2}\left(e^{2i\tan^{-1}\frac{D_r}{R}}\right)\Big\}\Big] \quad \text{(B.3)}$$

$$\sigma_{ic}^2 = \frac{D_r}{R}\Big[\frac{-i\pi^2}{6}+\pi\log 2-\frac{1}{D_r}\Big\{\tan^{-1}\frac{D_r}{R}\Big((-R-iD_r)\tan^{-1}\frac{D_r}{R}+2D_r\log\left(1-e^{2i\tan^{-1}\frac{D_r}{R}}\right)\Big) - iD_r\xi_{i_2}\left(e^{2i\tan^{-1}\frac{D_r}{R}}\right)\Big\}\Big] - \left(\tan^{-1}\frac{D_r}{R}-\frac{D_r}{R}\ln\left[\frac{D_r}{\sqrt{R^2+D_r^2}}\right]\right)^2 \quad \text{(B.4)}$$

[4] M. K. Ochi, *Ocean waves: the stochastic approach. Vol. 6*. Cambridge University Press, New York, USA, 2005.

[5] A. D. Jones, A. J. Duncan, A. L. Maggi, D. W. Bartel, and A. Zinoviev, "A detailed comparison between a small-slope model of acoustical scattering from a rough sea surface and stochastic modeling of the coherent surface loss," *IEEE J. Oceanic Eng.*, vol. 41, no. 3, pp. 689–708, 2015.

[6] Z. Zou and M. Badiey, "Effects of wind speed on shallow-water broadband acoustic transmission," *IEEE J. Oceanic Eng.*, vol. 43, no. 4, pp. 1187–1199, 2017.

[7] M. Naderi and M. Pätzold, "Design and analysis of a one-dimensional sea surface simulator using the sum-of-sinusoids principle," in *IEEE OCEANS, Washington, USA*, 2015, pp. 1–7.

[8] E. I. Thorsos, "The validity of the kirchhoff approximation for rough surface scattering using a Gaussian roughness spectrum," *J. Acoust. Soc. Am.*, vol. 83, no. 1, pp. 78–92, 1988.

[9] M. Rawat, B. Lall, and S. Srirangarajan, "Angle of arrival distribution in an underwater acoustic communication channel with incoherent scattering," *IEEE Access*, vol. 8, pp. 133 204–133 211, 2020.

[10] M. J. Soja, "Electromagnetic models of bistatic radar scattering from rough surfaces with Gaussian correlation function," Master's thesis, Department of Radio and Space Science, CTH, G¨oteborg, Sweden, 2009.

[11] P. H. Dahl, "High-frequency forward scattering from the sea surface: The characteristic scales of time and angle spreading," *IEEE J. Oceanic Eng.*, vol. 26, no. 1, pp. 141–151, 2001.

[12] R. Janaswamy, "Angle and time of arrival statistics for the Gaussian scatter density model," *IEEE Trans. Wireless Commun.*, vol. 1, no. 3, pp. 488–497, 2002.

[13] A. Abdi and H. Guo, "Signal correlation modeling in acoustic vector sensor arrays," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 892–903, 2008.

[14] J. Choi and H.-T. Choi, "Multi-target localization of underwater acoustic sources based on probabilistic estimation of direction angle," in *IEEE OCEANS, Genova, Italy*, 2015, pp. 1–6.

[15] Y. Zhang, Y. Li, Y. Zhang, and T. Jiang, "Underwater anchor-AUV localization geometries with an isogradient sound speed profile: A CRLB-based optimality analysis," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8228–8238, 2018.

[16] X. Cheng, Y. Li, C.-X. Wang, X. Yin, and D. W. Matolak, "A 3D geometry-based stochastic model for unmanned aerial vehicle MIMO ricean fading channels," *IEEE IoT J.*, vol. 7, no. 9, pp. 8674–8687, 2020.

[17] H. Jiang, Z. Zhang, C.-X. Wang, J. Zhang, J. Dang, L. Wu, and H. Zhang, "A novel 3D UAV channel model for A2G communication environments using AoD and AoA estimation algorithms," *IEEE Trans. Commun.*, vol. 68, no. 11, pp. 7232–7246, 2020.

[18] C. Huang, A. F. Molisch, Y.-A. Geng, R. He, B. Ai, and Z. Zhong, "Trajectory-joint clustering algorithm for time-varying channel modeling," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 1041–1045, 2019.

[19] E. Baktash, M. J. Dehghani, M. R. F. Nasab, and M. Karimi, "Shallow water acoustic channel modeling based on analytical second order statistics for moving transmitter/receiver," *IEEE Trans. Signal Process.*, vol. 63, no. 10, pp. 2533–2545, 2015.

[20] A. Amar, Y. Buchris, and M. Stojanovic, "Angle-of-arrival-based detection of underwater acoustic OFDM signals," in *IEEE SPAWC, Stockholm, Sweden*, 2015.

[21] M. R. Mousavi, M. Karimi, and M. Farrokhrooz, "Statistical properties of high frequency acoustic fluctuations induced by rough sea surface," in *IEEE OCEANS, Hampton Roads, VA, USA*, 2012, pp. 1–5.

[22] M. Rawat, B. Lall, and S. Srirangarajan, "Analysis of underwater acoustic channel based on angle of arrival distribution," in *IEEE COMSNETS, Bengaluru, India*, 2020, pp. 1–4.

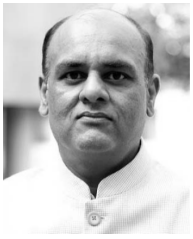
0018-9545 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 06,2021 at 06:38:40 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/TVT.2021.3131504, IEEE Transactions on Vehicular Technology

[23] P. H. Dahl, "On the spatial coherence and angular spreading of sound forward scattered from the sea surface: Measurements and interpretive model," *J. Acoust. Soc. Am.*, vol. 100, no. 2, pp. 748–758, 1996.
[24] H. Medwin and C. S. Clay, *Fundamentals of acoustical oceanography*. Academic press, USA, 1997.
[25] C. M. Bishop, *Pattern recognition and machine learning*. Springer, New York, USA, 2006.
[26] A. G. Zajic, "Statistical modeling of MIMO mobile-to-mobile underwater channels," *IEEE Trans. Veh. Technol.*, vol. 60, no. 4, pp. 1337–1351, 2011.
[27] M. Naderi, A. G. Zajić, and M. Pätzold, "A nonisovelocity geometry-based underwater acoustic channel model," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 2864–2879, 2017.
[28] E. I. Thorsos, "Surface forward scattering and reflection," (No. APL-UW-7-83), Applied Physics Laboratory, Washington University, Seattle, Tech. Rep., 1984.
[29] M. R. Dhanak and N. I. Xiros, *Springer handbook of ocean engineering*. Springer, 2016.
[30] K. Sarkar and A. Prosperetti, "Coherent and incoherent scattering by oceanic bubbles," *J. Acoust. Soc. Am.*, vol. 96, no. 1, pp. 332–341, 1994.
[31] D. Jackson and D. Tang, "The mutual scattering cross section," *The Journal of the Acoustical Society of America*, vol. 146, no. 6, pp. 4611–4624, 2019.
[32] A. Papoulis and H. Saunders, *Probability, random variables and stochastic processes, ed. 4*. Mc-Graw Hill, New York, USA, 1989.
[33] K. L. Williams, E. I. Thorsos, and W. Elam, "Examination of coherent surface reflection coefficient (CSRC) approximations in shallow water propagation," *J. Acoust. Soc. Am.*, vol. 116, no. 4, pp. 1975–1984, 2004.
[34] M. Evans, N. Hastings, and B. Peacock, "Triangular distribution," *Stat. Distrib.*, vol. 3, pp. 187–188, 2000.
[35] D. Bailey, P. Borwein, and S. Plouffe, "On the rapid computation of various polylogarithmic constants," *Math. Comput.*, vol. 66, no. 218, pp. 903–913, 1997.

Manishika Rawat (S'16) received B.Tech. degree in Electronics and Communication Engineering and M.Tech. degree in Digital Signal Processing from G. B. Pant Engineering College, Pauri Garhwal, Uttarakhand, India, in 2010 and 2012, respectively. She is currently pursuing Ph.D. from the Indian Institute of Technology Delhi, India.

Her research interests include Underwater acoustic communication, statistical channel modeling, cooperative techniques, angle of arrival distribution, coherent and incoherent scattering, Doppler effect, and vector sensor receivers.



Brejesh Lall (Member, IEEE) received the B.E. degree in electronics and communication engineering and the M.E. degree in signal processing from the Delhi College of Engineering, New Delhi, in 1991 and 1992, respectively, and the Ph.D. degree in signal processing from the Indian Institute of Technology (IIT) Delhi, New Delhi, in 1999.

He is a Professor with the Department of Electrical Engineering, Indian Institute of Technology Delhi. In 1997, he joined Aricent. He joined IIT Delhi as an Assistant Professor in 2005. He worked as an Associate Professor from 2010 to 2018. He is actively working in the area of image processing, signal processing, and communication. The areas include, object representation, tracking and classification, odometry, depth map generation, representation and rendering. He is also exploring vector sensor-based underwater acoustic communications, and performance issues in molecular communications.

Seshan Srirangarajan (S'05-M'08) received the B.E. degree from the University of Mumbai, India in 2001, and the M.S. and Ph.D. degrees from the University of Minnesota in 2005 and 2008, respectively, all in Electrical Engineering. He is currently an Assistant Professor in the Department of Electrical Engineering at the Indian Institute of Technology Delhi. Prior to this he was an IRCSET/Intel Postdoctoral Fellow with the Nimbus Centre for Embedded Systems Research at the Cork Institute of Technology, Ireland. From 2008 to 2011, he was a Research Fellow with the Intelligent Systems Center at the Nanyang Technological University, Singapore. In 2005–06, he was with the Wireless Technologies Group at Honeywell Technology Center, Minneapolis, Minnesota, as an intern. His research interests span the broad areas of signal processing and wireless communications with current focus being on positioning in wireless networks, sensor networks, compressed sensing, and array signal processing.

0018-9545 (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 December 06,2021 at 06:38:40 UTC from IEEE Xplore. Restrictions apply.

Stumpy-Paced Green Power Generation System

Ashok Kumar Singh[a], P. Kumar[b], R. Kishore[b], Phool Singh[c], Neeraj Kant[c], Pushpendra Singh[c], Gopal Nandan[d], A. Jaiswal[a], R. Prasad[b]

[a]*Mechanical Engineering Department, Galgotias College of Engineering and Technology, Greater Noida, UP, India* [b]*Applied Sciences Department, Galgotias College of Engineering and Technology, Greater Noida, UP, India*

[c]*Mechanical Engineering Department, Delhi Technological University, Delhi, India*

[d]*Mechanical Engineering Department, Amity University Uttar Pradesh, Noida, India*

[*]Corresponding author

agashok26@gmail.com*, pratap.kumar@galgotiacollege.edu, dr.rajeevkishorepandey@gmail.com, phoolsinghdce@gmail.com, neerajkant00007@gmail.com, Dr.psinghs@gmail.com, gnandan@amity.edu, megcet097@gmail.com, rajendra.prasad@galgotiacollege.edu

2021 4th International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE) | 978-1-6654-1429-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/RDCAPE52977.2021.9633633

Abstract: **In the present study, it has been emphasized that the specified gear mechanism or techno-suitable power transmission device may help to improve the power generation capabilities of the turbo-ventilators working at low-medium speeds. This set-up is indeed required to establish a clean and ventilated space at work-stations, theaters, assembly halls, stadiums, restraints, offices, etc. In this study, a generalized observation has been overviewed based on its development, and uses. Further, with a suggested utilization of gear mechanism with other associations along with a fair discussion and conscious conclusions have been given in the present work of approach. In this way, this communication builds a indeed gate way towards the utilization of freely available clean and green natural energy for saving the human beings and society in an eco-friendly way.**

Keywords: Roof top ventilator; Electric power; Green energy; Stumpy-paced.

I. INTRODUCTION

The industries, and other indoor spaces continuously producing hot air due to absorption of heat from heat libration electrical-electronic appliances, human bodies, etc. which goes out to the atmosphere due to lighter in weight and density in the replacement of fresh air coming through a confined passage of ventilators from the outside atmosphere. The freely out going hot air through a confined passage is further utilized to run the roof top turbo ventilators from one of the primary source and another one is air velocity available from open atmosphere as a natural renewable sources. Whereas, other natural resources as solar energy is also helpful to boost up the performance of renewable energy systems, along with the solar heat receivers [1] and related other heat receiving techniques [2, 3] like solar water heat treatment [4, 5], steam production [6], electrical power production [7], under the influence of ecological patterns [8, 9] energy matrix [10, 11] with economical observation and perspectives [12, 13]. In these, the renewable energy utilizations are involved for power production without any environmental harm. Further, India installed power production system through wind renewable energy in February 2021, which is the 4th biggest renewable power generation plant [14]. This project utilizes the available wind energy at variable wind velocities to run the system. Also, this wind operated ventilators are efficiently working for different wind directions (horizontal, axial, radial, and vertical) for the production of electric power [15]. As, these wind ventilators are required to operate with sufficient wind velocities in an arid zone only as represented in Fig. 1.

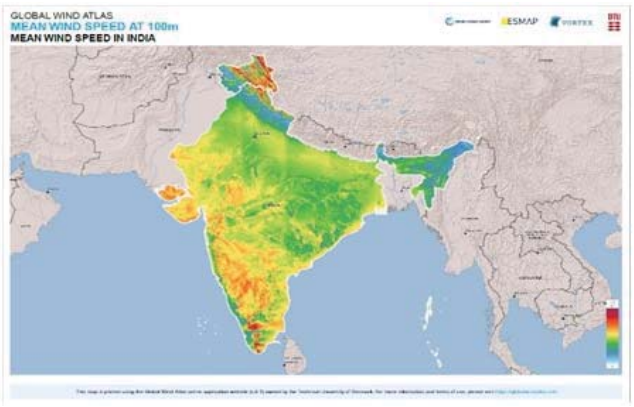


Fig. 1. Average wind velocity in Indian continent

So, it is better to use turbo-ventilators working at low-medium a speed which establishes a clean and ventilated space at work-stations, theaters, assembly halls, stadiums, restraints, offices, etc. and further, utilized for the production of electrical power also [16]. The wind operated roof top turbo ventilators are effective for high regions, arid areas for which low-medium speed wind can operate the mechanism of these ventilators to establish a clean and ventilated space at work-stations. These set up utilizes spherically designed vents made up of metals (erosion resistant, corrosion resistant with low thermal expansion coefficients). The roof top installed turbo wind ventilators are rotating with the effect of hot air coming from indoor spaces, electrical, electronic appliances, human bodies, and related objects, etc. and this rotates the turbo wind ventilators. As a result, a wind siphon effect developed that drops the air pressure significantly into the working space of work-stations, theaters, assembly halls, stadiums, restraints, offices, etc. Finally, the wind siphon effect and air pressure drop cumulatively sucks the fresh air from outside atmosphere through fresh air vent passages [17]. As per the global power generation which is mainly dependent on fossil fuels (>60%) as shown in Fig. 2. Now in the recent years, consumption of renewable energy sources are increased (>8%) globally. So the carbon savings and environmental protection are further keep better as shown in Fig. 3.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:04:22 UTC from IEEE Xplore. Restrictions apply.

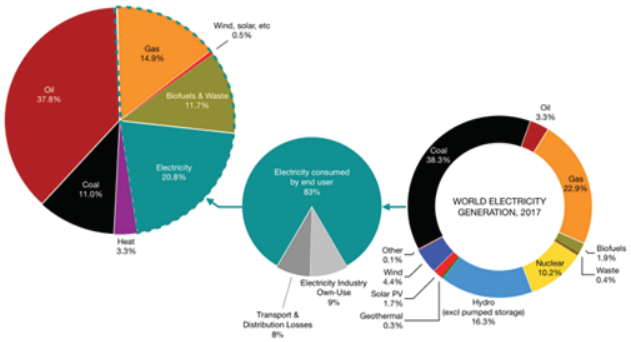


Fig. 2. Global energy facts [16]

The fundamental objectives of the present study are focusing the following frame of scheme as written herewith.

a) To study the development in the field of roof top installed wind ventilator fans and related technologies inbuilt within and the regions for which, these techniques are quite suitable for the applications.
b) To find the opportunities for low wind or low arid areas under the mono directional use of these roof top installed wind ventilator

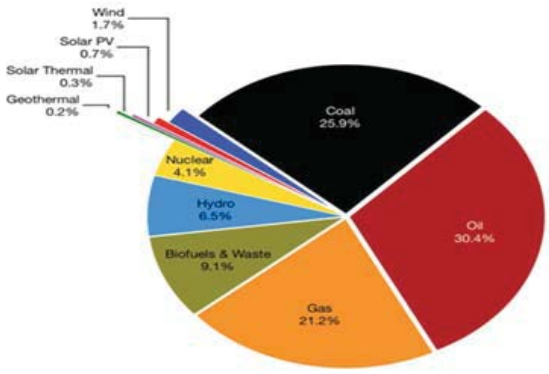


Fig. 3. Bifurcation of different energy possessions [16]

c) To apply sophisticated additional gear mechanism to improve the power output of the overall establishment (roof top installed wind ventilator) under the low speed wind capacity by increasing the rotation of wind ventilator metallic vanes.

II. CONSTRUCTIONAL DETAILS

It has been observed that the specified gear mechanism or techno-suitable power transmission device may help to improve the power generation capabilities of the turbo-ventilators working at low-medium speeds. And, such set-ups are required to establish a clean and ventilated space at work-stations, theaters, assembly halls, stadiums, restraints, offices, etc. under the wind potential capabilities. The wind ventilator regulates the warm air of occupied space from fresh air of atmospheric air through the air vent passages (metallic) as represented in Figs. 4(a)-4(b). As the hot air (heated up by accumulating heat coming out from various heat releasing sources available inside the workstation) is lighter in weight due to lowered density than the colder fresh air and this causes dual effect to run the roof top ventilators as mentioned,

a) The development of pressure drop at the ventilator vanes due to the upside movement of hot air.
b) Extra push developed by the fresh air over the hot air that provides additional up-lift thrust force and corresponding velocity to the hot air.

In this way, roof top turbo ventilators are continuously spinning throughout the day due to air exchange circulations.

In the development of turbo ventilators, computational fluid dynamic analyses is one of the important tool to simulate the air flow and heat effect for the variable vent (throat) diameter which had been taken as 300mm for greater circulation speed under given air flow [18]. Further, axial flux type permanent magnet (AFPM) as shown in Fig. 5 and Table 1 which is utilized to generate the corresponding force in between the armature winding, and during the rotation of armature, direct current (DC) is induced that can be utilized in a controlled manner to run the suitable electrical appliances [21]. All these combination with a wind turbine make the roof top turbo ventilators as a mini power plant [19]. Also, the DC power is converted accordingly based on the suitability of the available electrical appliances for the sound use of those [20]. The developed power can be utilized to improve the operating conditions and performance of other renewable energy systems, as additional benefits of such stumpy-paced green power generation systems for the particular solar operating devices such as solar desaltification systems utilized to treat the brackish water [22].

Table 1 Variation in speed of rotation against the wind velocities along with the effect or without the effect for AFPM base gears mechanism [21].

Wind velocity (m/s)	Speed of Rotation (rpm)	AFPM Rotational speed without gear (rpm)	AFPM Rotational speed with gear (rpm)
1.0	30	30	60
2.0	45	45	90
2.75	56	56	112
3.44	62	62	124
4.55	107	107	214
5.25	120	120	240
7.60	132	132	264
8.87	151	151	302
9.5	195	195	390

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:04:22 UTC from IEEE Xplore. Restrictions apply.

Fig. 4 (a). Roof top ventilator

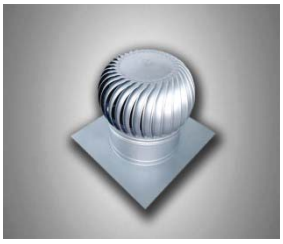


Fig. 4 (b). Roof top ventilator

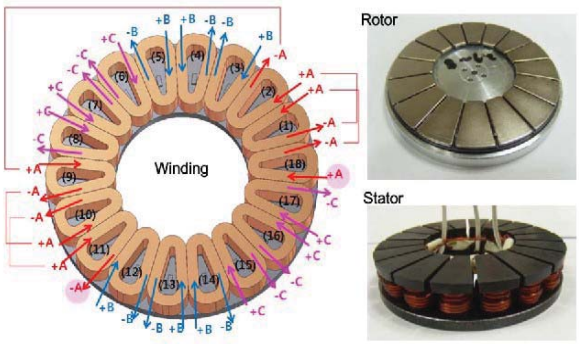


Fig. 5. Axial flux type permanent magnet generator and rectifier [21]

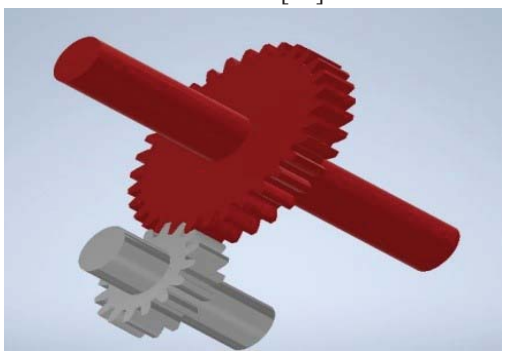


Fig. 6. Three dimensional model of gear mechanism

Also, a sophisticated gear mechanism (shown in Fig. 6) with the detailed specification and design parameters as given in Table 2 for the proper transmission of generated mechanical power towards the electric power generator. All these components are sequentially assembled as shown in Fig. 7, which makes the complete establishment as a roof top turbo ventilator with a proficient power transmission system (gear mechanism) and power generator.

Table 2 Gear mechanism association and design parameters utilized for a particular system

Design parameters	**Values**
Module (mm)	2
Pinion teeth	16
Gear teeth	32
Gear ratio	2
Pinion-gear center (mm)	48
Circular pitch of pinion (mm)	32
Circular pitch gear (mm)	64
Addendum (mm)	2
Dedendum (mm)	2.5
Tooth thickness (mm)	3.2
Bottom clearance (mm)	0.5
Whole depth (mm)	4.5

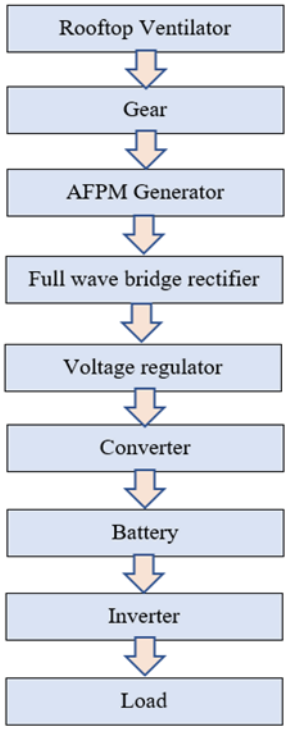


Fig. 7. Probable layout of the assembly

III. WORKING

With the help of a specified gear mechanism or techno-suitable power transmission device, one can improve the power generation capabilities of the turbo-ventilators working at low-medium speeds. And, such set-ups are required to establish a clean and ventilated space at work-stations, theaters, assembly halls, stadiums, restraints, offices, etc. under the wind potential capabilities. The wind ventilator regulates the warm air of occupied space from fresh air of atmospheric air through the air vent passages. As the hot air (heated up by accumulating heat coming out from various heat releasing sources available inside the workstation) is lighter in weight due to lowered density than the colder fresh air with the dual effect to run the roof top ventilators by developing the pressure drop at the ventilator vanes due to the upside movement of hot air and extra push offered by the fresh air over the hot air that provides additional up-lift thrust force and corresponding velocity to the hot air. In this way, roof top turbo ventilators are continuously spinning throughout the day due to air exchange circulations. In the development of turbo ventilators with better effect of air flow, the variable vent (throat) diameter may be taken as 300mm for greater circulation speed under given air flow [18]. Further, axial flux type permanent magnet (AFPM) which is utilized to generate the corresponding force in between the armature winding, and during the rotation of armature, direct current (DC) is induced that can be utilized in a controlled manner to run the suitable electrical appliances. All these combination with a wind turbine make the roof top turbo ventilators as a mini power plant. Also, the DC power is converted accordingly based on

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:04:22 UTC from IEEE Xplore. Restrictions apply.

the suitability of the available electrical appliances for the sound use of those.

IV. DISCUSSION

It is quite necessary to have a sound association in between the integrated components of the roof top turbo ventilators with gear mechanism and electric power generator for the development of greater amount of electricity with least losses (mechanical or electrical). The gear mechanism is one of the most important linking elements which may be further tested on real background for its feasible viability and potential of utilization during the service of turbo ventilators.

Hence, the proper utilization of green energy sources along with the renewable energy sources of nature accumulatively builds an eco-friendly mechanism and sustainable technology for the production of a substantial amount of electric power. These electric powers can be utilized to run electrical appliances as fan, LED bulbs, blowers, coolers, etc. by converting it into alternating current power where ever necessary.

V. CONCLUSION

In the current study, the generalized idea of working, construction, and the developments of roof top turbo ventilators have been revealed. The conscious conclusions of the present study have been framed as written here under.

a) Sophisticated association is between the integral parts is utmost required to produce greater electric power with least thermal, mechanical, and electrical losses.
b) Gear mechanism is one of the important mechanical power transmission element, which requires on line testing for it proper and efficient functioning, that ensures smooth working of he system as well as transmits greater mechanical power to the generator.
c) Roof top turbo ventilators are quite successful for the domestic purposes also, and should be utilized by modifying the gear mechanism assembly along with the ventilator and generator system.

REFERENCES

[1] A.K. Singh, Samsher, Analytical study of evacuated annulus tube collector assisted solar desaltification system: A review, Solar Energy 207 (2020) 1404–1426.

[2] A.K. Singh, D.B. Singh, A. Mallick, Harender, S.K. Sharma, N. Kumar, V.K. Dwivedi, Performance analysis of specially designed single basin passive solar distillers incorporated with novel solar desalting stills: A review, Solar Energy, 185 (2019) 146–164.

[3] A.K. Singh, D.B. Singh, V.K. Dwivedi, N. Kumar, J.K. Yadav, A Review of Performance Enhancement in Solar Desalination Systems with the Application of Nanofluids, ICACCCN, (2018) 814 – 819.

[4] A.K. Singh, D.B. Singh, V.K. Dwivedi, G.N. Tiwari, A. Gupta, Water purification using solar still with/without nano-fluid: A review, Materials Today: Proceedings 21 (2020) 1700–1706.

[5] A.K. Singh, Samsher, A Review Study of Solar Desalting Units with Evacuated Tube Collectors, Journal of Cleaner Production 279 (2021) 123542.

[6] A.K. Singh, An inclusive study on new conceptual designs of passive solar desalting systems, Heliyon 7 (2020) e05793.

[7] A.K. Singh, R.K. Yadav, D. Mishra, R. Prasad, L.K. Gupta, P. Kumar, Active solar distillation technology: A wide overview, Desalination 493 (2020) 114652.

[8] A.K. Singh, B.K. Srivastava, V. Shankar, R. Prasad, R.K. Yadav, A.K. Shukla, G. Nandan, P. Kumar, Energy matrices and enviroeconomical perspective of solar decontamination system, Materials Today: Proceedings, (2020). https://doi.org/10.1016/j.matpr.2020.09.198

[9] A.K. Singh, D. Singh, M.K. Lohumi, B.K. Srivastava, H.P. Gupta, R. Prasad, Potable Water Production by Single Slope Active Solar Distillation Unit-A Review. Advances in Energy and Environment. Lecture Notes in Civil Engineering, Springer 142 (2021) 31-42.

[10] A.K. Singh, D.B. Singh, A. Mallick, N. Kumar, Energy matrices and efficiency analyses of solar distiller units: A review. Solar Energy, 173 (2018) 53 – 75.

[11] A.K. Singh, Samsher, Material conscious energy matrix and enviro-economic analysis of passive ETC solar still, Materials Today: Proceedings 38 (2021) 1-5.

[12] A.K. Singh, G. Dwivedi, B.K. Srivastava, B.K. Tiwari, P. Kumar, R.K. Yadav, M. Singh, A.K. Shukla, G. Nandan, Study of analytical observations on energy matrix for solar stills, Materials Today: Proceedings, (2021). https://doi.org/10.1016/j.matpr.2021.02.109

[13] A.K. Singh, D. Singh, M.K. Lohumi, B.K. Srivastava, H.P. Gupta, R. Prasad, Potable Water Production by Single Slope Active Solar Distillation Unit-A Review. Advances in Energy and Environment. Lecture Notes in Civil Engineering, Springer 142 (2021) 31-42.

[14] A. Kumar, K. Kumar, N. Kaushik, S. Sharma, S. Mishra, Renewable energy in India: Current status and future potentials, Renewable and Sustainable Energy Reviews 14 (8) (2010) 2434-2442.

[15] R. Shinde, V. Lavhale, A. Nair, S. Pawar, R. Mahajan, Generation of Electric Power using Turbo Ventilators,

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:04:22 UTC from IEEE Xplore. Restrictions apply.

International Journal of Current Engineering and Technology 6 (2016) 89 – 93.

[16] G.K. Jadhav, P.M. Ghanegaonkar, Performance Characteristics of Turbo Ventilator: A Review, International journal of innovative technology and research volume 3 (4) (2015) 2276 – 2282.

[17] K. Dongre, A.A. Meshram, A.R. Bhoyar, R.U. Tamgade, A.R. Kapse, S. Jethani, Electricity generation by using rooftop ventilator, International Research Journal of Engineering and Technology 7(2) (2020) 3006-3008.

[18] G.K. Jadhav, P.M. Ghanegaonkar, S. Garg, Experimental and CFD Analysis of Turbo Ventilator, Journal of Building Engineering 6 (2016) 196-202.

[19] S. Dangeam, An electric generator is driven by a roof ventilator, Energy Procedia 9 (2011) 147-158.

[20] D.C. Shatri, M. Irwanto, A.N. Syafawati, S.S. Shema, Power generation roof ventilator, International Conference on Environment and Industrial Innovation 12 (2011) 183-187.

[21] M.R. Minaz, M. Celebi, Design and analysis of a new coreless PMSG with three rotor and double stator, Results in Physics 7 (2017) 183-188.

[22] A.K. Singh, Samsher, Tech-En-Econ-Energy-Exergy-Matrix (T4EM) Observations of Evacuated Solar Tube Collector Augmented Solar Desaltification Unit: A Modified Design Loom, Materials Today: Proceedings, (2021). https://doi.org/10.1016/j.matpr.2021.09.088

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 22,2021 at 07:04:22 UTC from IEEE Xplore. Restrictions apply.

THE JOURNAL OF PHYSICAL CHEMISTRY C
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCC

Article

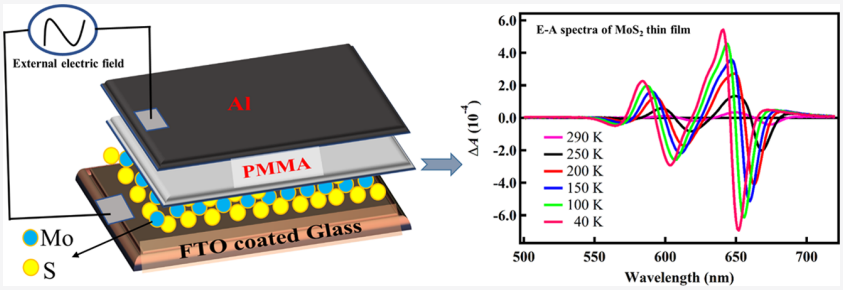
Temperature-Dependent Electric Field-Induced Optical Transitions of 2D Molybdenum Disulfide (MoS_2) Thin Films: Temperature-Dependent Electroabsorption and Absorption

Mohan Singh Mehata,*[,‖] Mrityunjay Kumar Singh,[‖] Kamlesh Awasthi, Prateek Sharma, Shailesh Rana, and Nobuhiro Ohta*

Cite This: https://doi.org/10.1021/acs.jpcc.1c06706

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Two-dimensional (2D) layered MoS_2 nanosheets (NSs) possess many unique properties and hold great potential for various applications. Herein, MoS_2 NSs were synthesized by a hydrothermal method. The as-synthesized MoS_2 NSs are crystalline and layered. Absorption and electroabsorption (E-A) spectra of MoS_2 doped in a poly(methyl methacrylate) (PMMA) thin film were measured at different temperatures (290−40 K). The E-A spectra detected at the second harmonic of the modulation frequency of the applied electric field were analyzed using an integral method by considering the Stark effect as a dominant feature. The absorption spectra consist of seven transitions, among which five transitions are contributed to the E-A spectra. It is found that the changes in the electric dipole moment and polarizability of each transition determined at different temperatures increase substantially with decreasing temperature. Electronic resonance states identified for low-energy excitonic bands of MoS_2 NSs showed prominence E-A signals. The study is essential to understand the electronic structure in the photoexcited state, which is important for applications of MoS_2 NSs to optoelectronic devices.

INTRODUCTION

Molybdenum disulfide (MoS_2) is a two-dimensional (2D) transition metal dichalcogenide formed with the atomic layer of molybdenum sandwiched between two sulfur atom layers. In MoS_2 nanosheets (NSs), there is a strong intralayer covalent bonding of Mo atoms with S atoms in a trigonal prismatic arrangement. A bulk MoS_2 crystal is an indirect band gap semiconductor material with several S−Mo−S layers and switches to a direct band gap when it is brought to a monolayer. When the film exfoliates from layers to a single layer, it has a tunable band gap,[1,2] and various catalytic and sensing applications are possible. For example, MoS_2 is used as a photocatalyst because of its unique optical band gap, which falls in the visible region and overcomes the shortcoming of other renowned photocatalysts (TiO_2, ZnO, SnO_2, etc.).[3−9]

MoS_2 monolayers are also applicable to optoelectronic devices such as phototransistors,[10,11] where light would be directly converted into current, and photovoltaic solar cells,[12] i.e., by the formation of a heterojunction with p-Si. To design and develop more efficient optoelectronic devices, including phototransistors or photovoltaic cells with the MoS_2 film, a deep understanding of the electronic structure and dynamics in the excited states of the MoS_2 film reached by photoexcitation is inevitable. In a single crystal of MoS_2, strong absorption bands (A and B) are observed in the 700−550 nm region, and two other much stronger bands (C and D) are observed in the 500−350 nm region.[13,14] The former two bands are interpreted in terms of Mott−Wannier delocalized exciton absorption, resulting from the direct transition from the highest spin-split valence band to the lowest conduction band, while the latter two bands, which show only a slight energy shift with temperature, are attributed to the transitions from low lying levels in the valence band.[15,16] These four strong

Received: July 28, 2021
Revised: November 16, 2021

Downloaded via DELHI TECHNOLOGICAL UNIV on December 6, 2021 at 06:31:41 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

ACS Publications
© XXXX American Chemical Society

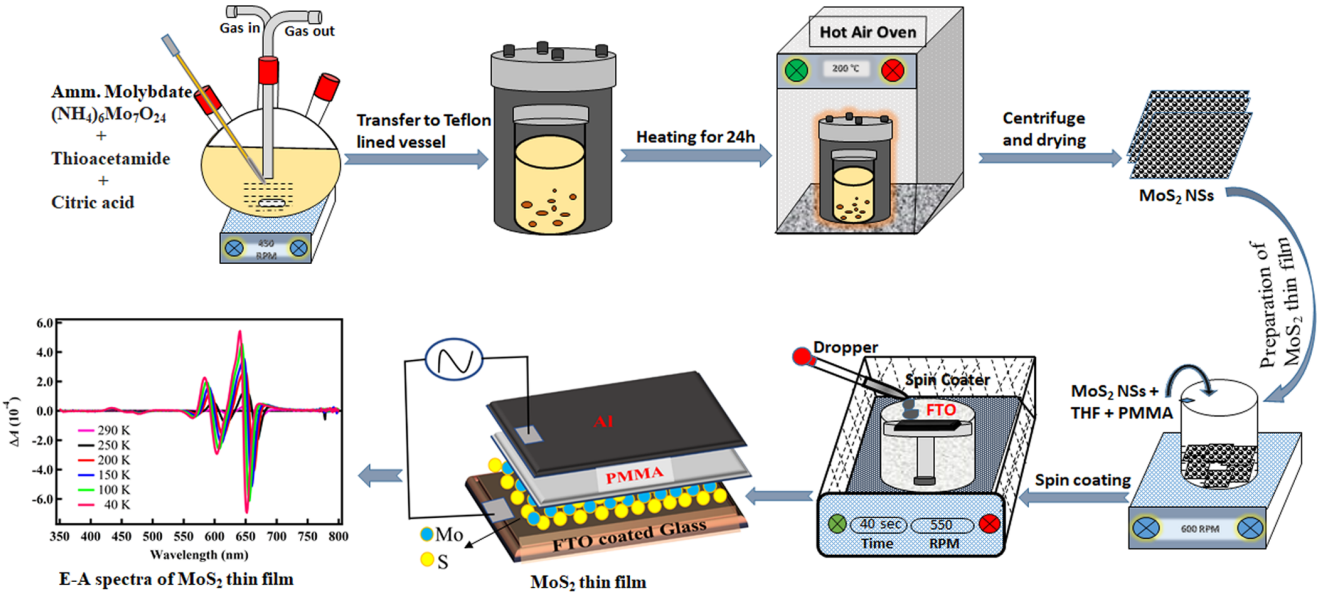


Figure 1. Schematic representation of the synthesis of the MoS_2 NS thin film.

bands are clearly observed in the present samples of MoS_2 nanosheets, as shown in the results. The photocurrent of a single crystal of MoS_2 measured with excitation in the range of 800−400 nm indicates that current efficiency decreases with increasing excitation energy, i.e., decreasing excitation wavelength, and depends on temperature.[13,14] However, the electronic characteristic of each absorption band is not known well. To obtain detailed information about the electronic structure and dynamics in the excited states, an electroabsorption (E-A) study, i.e., measurements of E-A spectra, is very useful because the change in electric dipole moment and polarizability can be determined from the analysis of the E-A spectra.[17−19] Based on the results, the charge-separated character and distribution of charge density in the photoexcited state can be discussed.[20−25]

Another advantage of the measurements of E-A spectra is that the applied electric field could enhance the intensity of hidden or inherently weak absorption bands, depending on the magnitude of the change in electric dipole moment and/or polarizability following photoexcitation.[26,27] In the MoS_2 crystal, many weak exciton bands were pointed out to be located in the shorter wavelength region of A and B bands.[14] Besides the exciton absorption caused by the transition from the neutral ground state to a neutral exciton state, absorption bands that can be assigned to the absorption of a trion, i.e., a quasiparticle composed of two electrons and a hole, were also demonstrated.[28] By analyzing the E-A spectra, therefore, there may be a possibility that absorption bands that cannot be realized in usual absorption spectra are clearly identified.

Bordas and Davis[29] measured the E-A spectra of a single crystal of MoS_2 observed at the same modulation frequency as the applied voltage at room temperature in the region of ∼695−445 nm (from ∼1.8 to 2.8 eV), i.e., in the region of A, B, C, and D bands. Detailed analysis was not done, but it has been mentioned that the shape and the size of the spectrum depended on the direction of the applied voltage and that band A showed a shift along with slight broadening in the presence of the applied electric field. Vella et al.[30] observed the E-A spectra of the MoS_2 monolayer at room temperature in the range of 800−520 nm by fabricating a transistor with indium tin oxide (ITO) as the transparent back-gate electrode and Ti/Au as a source and drain electrodes. They reported that the broadening of the absorption band, which occurred in the presence of the applied electric field, increased linearly with the applied field strength, contrary to known variants of the quadratic Stark effect, but the magnitude of the change in electric dipole moment or polarizability following photoexcitation was not mentioned. The E-A spectra of the MoS_2 monolayer in the region of A and B bands were also measured at room temperature with a strong terahertz (THz) field.[31] In this case, A and B bands were reported to be broadened by an electric field, and the linewidth was shown to depend on the square of the THz field. Thus, both linear and quadratic field effects may exist in MoS_2 films.

In the present study, we prepared MoS_2 NSs using a hydrothermal method and investigated the electric field effects on absorption of MoS_2 NSs in a poly(methyl methacrylate) (PMMA) film at various temperatures from 290 to 40 K. The E-A spectra detected at the second harmonic of the modulation frequency of the applied electric field confirms that lower energy exciton bands, i.e., A and B bands, split into $(\bar{A})$, (A^0) and $(\bar{B})$, (B^0) resonances, respectively, and show a different response from each other to the application of the external electric field. By considering the Stark effect as a dominant feature, the change in polarizability $(\Delta\bar{\alpha})$ and electric dipole moment $(|\Delta\mu|)$ following photoexcitation has been determined for each transition of resonance bands and higher energy exciton bands.

EXPERIMENTAL SECTION

All reagents were of AR (analytically pure) grade and received from Sigma-Aldrich. Ammonium heptamolybdate tetrahydrate was used as a precursor and primary source of Mo. Thioacetamide was taken as a sulfur source for the synthesis of the required compound. Also, citric acid was used to reduce reagents. PMMA obtained from Sigma-Aldrich was purified before a MoS_2 NSs-doped PMMA thin film was prepared with tetrahydrofuran (THF) as a solvent. Fluorine-doped tin oxide

https://doi.org/10.1021/acs.jpcc.1c06706
J. Phys. Chem. C XXXX, XXX, XXX−XXX

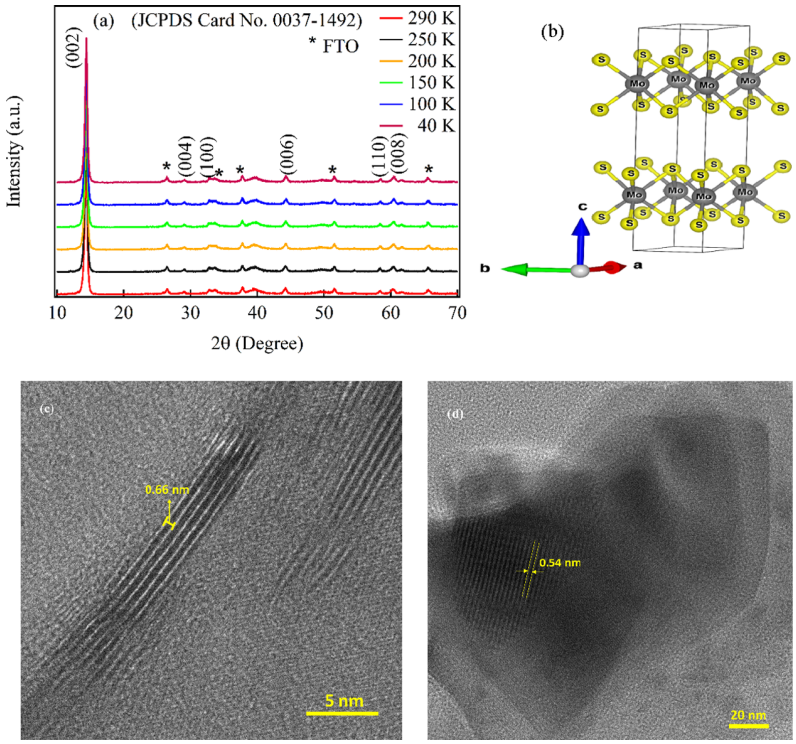


Figure 2. XRD patterns of MoS_2 NSs obtained at varying temperatures (a), hexagonal crystal structure (b), and HRTEM images with lattice fringes (c, d).

(FTO)-coated glass substrates were used for E-A measurements.

MoS_2 NSs were synthesized by dissolving 0.5 g of ammonium molybdate as a Mo source in deionized (DI) water with 0.2 g of citric acid for reducing Mo from Mo^{+6} to the Mo^{+4} state. The above mixture was allowed to stir for 2 h at room temperature, and afterward, 0.5 g of thioacetamide was added to the mixture as a sulfur source. Then, the mixture was transferred to a Teflon line autoclave at 200 °C for 24 h and cooled down to room temperature naturally. Finally, the obtained black precipitates were washed at 8000 rpm for 30 min with deionized water and ethanol. The black powder sample was kept in a hot air oven at 60 °C for 12 h for drying, and the final product was obtained as MoS_2 NSs.[32]

To prepare the solid thin film of MoS_2 NSs for the application of the external electric field, i.e., for the measurement of E-A spectra, a certain amount of MoS_2 NSs and PMMA were mixed in a selected volume of THF. After making a homogeneous mixture, it was spin-coated at fixed rpm on a FTO-coated glass plate. The FTO was pre-etched for about a quarter of its length before spin coating. The thickness of the MoS_2-doped thin film on the FTO substrate came out to be typically 0.42 μm, as recorded by an α step profiler (Veeco Dektak 150). The thickness of the film was used to assess the applied field strength (F) by dividing the applied voltage. The FTO layer served as an electrode. For the other electrode, an aluminum (Al) film having a thickness of about 20 nm was deposited on the MoS_2 thin film with a thermal evaporation technique. The semitransparent Al film and the FTO layer served as electrodes for the E-A measurements. The schematic representation of the synthesis of MoS_2 NSs and thin-film preparation is shown in Figure 1.

All of the steady-state measurements were recorded under vacuum conditions. The absorption spectra of MoS_2 NSs doped in a PMMA thin film were measured with a spectrophotometer (JASCO, FP-777). For the E-A measurements, a modulated AC voltage having a 1 kHz frequency was applied to the sample film, which was simultaneously shined by a 150 W continuous xenon lamp (from FP-777), incorporated, and controlled altogether in the apparatus, as described elsewhere.[20] The E-A signal, that is, the field-induced change in absorbance, i.e., ΔA, was determined from $-\frac{2\sqrt{2}\,\Delta I}{\ln 10 I}$, where I and ΔI are transmitted light intensity for excitation and its field-induced change, respectively. In the absorption and E-A measurements, the incident light was irradiated to an area of ~3 mm × 3 mm of the sample film, and the deviation in sample thickness in this area was ~5%. The observed absorption and E-A signal are regarded as the average one of MoS_2 film located in this area. The temperature-dependent E-A spectra were obtained using a V202C5LR cryostat, Daikin equipped with quartz optical windows. A temperature controller (Scientific Instruments Model 9600) was used to control the sample film temperature with a thermometer (silicon diode) attached to the sample. An X-ray diffractometer (Bede D1 with Cu Kα radiation) was used to measure the X-ray diffraction (XRD) patterns of the MoS_2 thin film at an acceleration voltage of 40 kV and an applied current of 50 mA. High-resolution transmission electron microscopy (HRTEM) images were recorded using a TALOS (Thermo Scientific Instruments) with an accelerating voltage of 80−200 kV.

RESULTS AND DISCUSSION

Figure 2a shows the XRD pattern of the as-synthesized MoS_2 NSs thin film observed in the range of 10−70° at a varying temperature range of 290−40 K. The XRD pattern exhibits the

https://doi.org/10.1021/acs.jpcc.1c06706
J. Phys. Chem. C XXXX, XXX, XXX−XXX

hexagonal crystal structure of MoS_2, as demonstrated in Figure 2b. No other impure phase was observed in the sample, indicating the formation of a pure hexagonal phase of MoS_2. The dominant peak of the MoS_2 crystal at 2θ = 14.28° designates the MoS_2 layers piled orderly along the (002) direction with a d-spacing of 0.61 nm. The diffraction peaks of MoS_2 at 2θ = 14.28, 44.21, and 60.31° show layered stacks along the (002), (006), and (008) planes, respectively. The (100) reflection at 32.58° and (110) reflection at 58.18° were observed in the as-synthesized samples. As the temperature reduced from 290 to 40 K, there is no variation in the peak intensity and peak position of the XRD pattern of MoS_2 NSs, indicating no change in the crystalline structure even at 40 K. The HRTEM images of the MoS_2 NS thin film, as shown in Figure 2c, prove the layered structure of MoS_2 up to a few layers, which are well stacked with an interlayer separation distance of 0.66 nm. Actually, each MoS_2 NS is considered to have five or six layers, resulting in a thickness of 3.3−3.9 nm. Figure 2d shows that the ordered lattice fringes of MoS_2 NSs with a d-spacing of 0.54 nm relates to the (002) oriented crystal plane of the 2H phase of MoS_2.

Figure 3a shows the absorption spectra of the as-synthesized MoS_2 NSs doped in a thin PMMA film in the range of 300−800 nm at various temperatures ranging from 290 to 40 K. At 290 K, four characteristic absorption peaks appeared and are designated as A, B, C, and D excitonic bands (Figure 3a). These bands correspond to A, B, C, and D bands reported in a crystal of MoS_2.[13,14] Exciton peaks A and B at 682 and 634 nm at 290 K, respectively, originated due to the spin−orbit splitting of interband excitonic transitions at the K point of the 2D Brillouin zone.[33,34] Absorption peaks C and D at 469 and 405 nm at 290 K, respectively, arise from the interband transitions between the higher density of state regions, from the occupied d_{z^2} orbital to the unoccupied d_{xy}, $d_{x^2-y^2}$ and d_{xz}, d_{yz} orbitals.[15,16] The broadening in overall absorption bands A, B, C, and D, which show the characteristic peaks of MoS_2 NSs, may be due to the Mie scattering-induced background.[35] The presence of direct gap transition and indirect gap transition may also result in spectral broadening. The monolayer is a direct gap semiconductor, and an indirect gap appears with the addition of the second layer. As more layers are added, the indirect gap becomes smaller, while the change in the direct gap is very small.[36] When MoS_2 NSs have five or six layers, the difference between the direct gap and the indirect gap may be estimated to be ∼0.7 eV, which may give a weak absorption tail in the longer wavelength region of the direct band gap transition. The nonuniform size distribution of the sheets may also contribute to the broadening of the individual peaks.

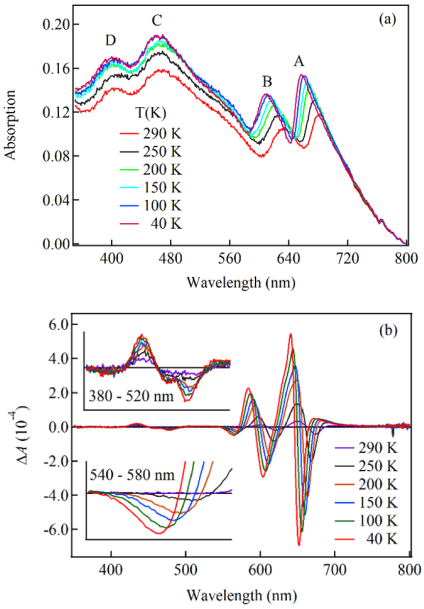


Figure 3. Absorption spectra (a) and E-A spectra (b) of the MoS_2 thin film obtained at different temperatures. The inset figure in (b) shows an extended view of the E-A spectra in ranges of 380−520 and 540−580 nm.

When the temperatures decreased from 290 to 40 K, absorption bands A and B showed a blue shift, while the spectral shift was negligible for C and D bands, as shown in Figure 3. The blue shifts with decreasing temperature could be due to a change in the orbital overlap, forming the band in thermally expanded crystals.[37] Such transformation is considered very small for C and D bands (Figure 3a).

Figure 3b shows the E-A spectra of the as-synthesized MoS_2 NSs doped in a thin PMMA film measured at the second harmonic of the modulation frequency of the applied electric field (1.0 kHz) at different temperatures (290−40 K) with a field strength of 0.5 MV cm^{-1}. The E-A signal is strong for the excitonic resonance bands (A and B) at room temperature. The E-A signal intensity increased and the spectrum shifted toward the higher energy side as the temperature decreased. In contrast, no shift in the E-A spectrum was observed in the short wavelength region corresponding to C and D bands, as shown in the inset of Figure 3b (extended view). A similar blue shift with decreasing temperature for A and B bands was also noticed in the absorption spectrum, as mentioned above. As the temperature dropped from 290 to 40 K, a new band appeared between 540 and 580 nm (17 241 to 18 518 cm^{-1}) and shifted toward the higher energy region, as selectively shown in the inset Figure 3b. The blue shift of A and B bands in the absorption and E-A spectra might be due to the lattice contraction in nanosheets.

The E-A spectra of MoS_2 NSs at 290 and 40 K were measured with different applied field strengths in the range of 0.1−0.5 MV cm^{-1}, and the results are shown in Figure 4a,b, respectively. The E-A spectra exhibit maxima at 652 nm (15 337 cm^{-1}) and 641 nm (15 601 cm^{-1}) and minima at 673 nm (14 859 cm^{-1}) and 652 nm (15 347 cm^{-1}) at 290 and 40 K, respectively. The E-A signals at two different temperatures are proportional to the square of the applied electric field (Figure 4c,d), indicating that the E-A signals of the excitonic bands originated from the quadratic Stark effect. It is noted that the E-A signals observed by Bordas and Davis[29] and Vella et al.[30] are regarded as linearly proportional to the applied electric field, in contrast to the present results.

Here, the Stark effect is considered to be a dominant feature. When an external electric field is applied to the MoS_2 NSs, the energy levels shift to the lower or higher side following the factors $-\Delta\mu F - \Delta\alpha(F^2)/2$, where $\Delta\mu$ and $\Delta\alpha$, respectively, are the differences in electric dipole moment and polarizability between the states concerned, namely, the electronically excited state (e) and the ground state (g), i.e., $\Delta\mu = \mu_e - \mu_g$ and $\Delta\alpha = \alpha_e - \alpha_g$. The electric field-induced change in the absorption spectrum of MoS_2 NSs at the wavenumber (ν) recorded at the second harmonic of modulation frequency, that is, $\Delta A(\nu)$, is taken as a sum of the zeroth, first, and second

https://doi.org/10.1021/acs.jpcc.1c06706
J. Phys. Chem. C XXXX, XXX, XXX−XXX

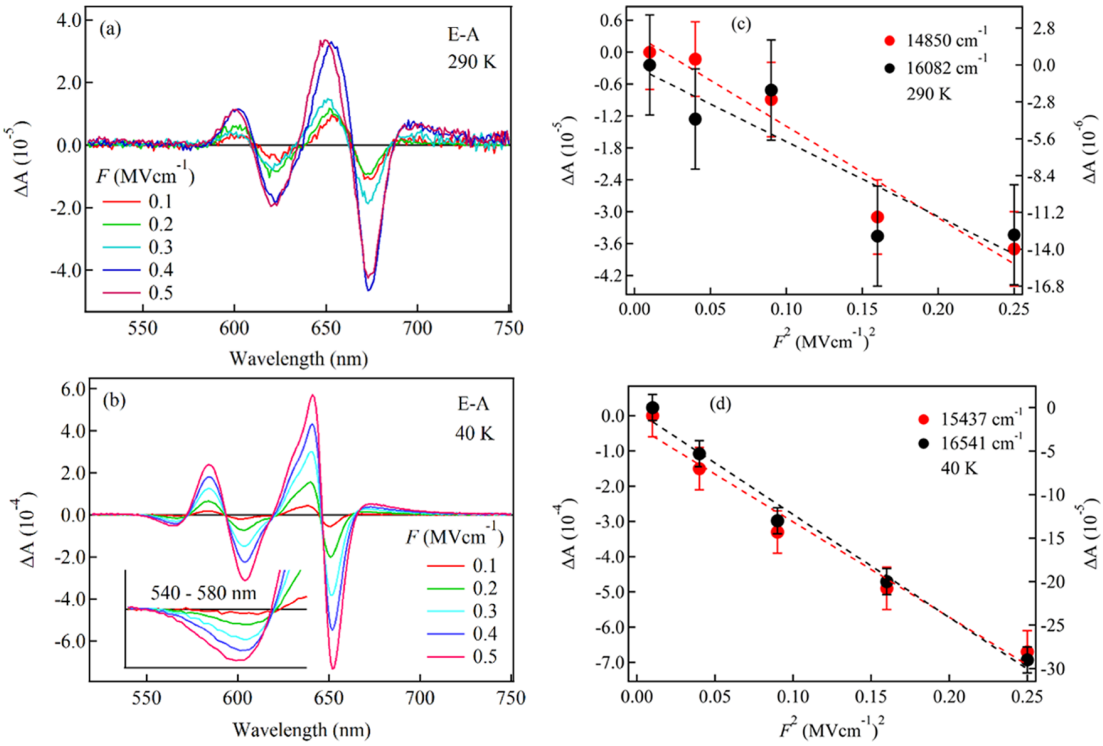


Figure 4. E-A spectra of the MoS_2 thin film observed with different electric field strengths at 290 K (a) and 40 K (b). Plots of the dependence of the applied field strength of ΔA as a function of squared field at 290 K (c) and 40 K (d). The inset figure in (b) shows an extended view of the E-A spectra at 40 K from 540 to 580 nm.

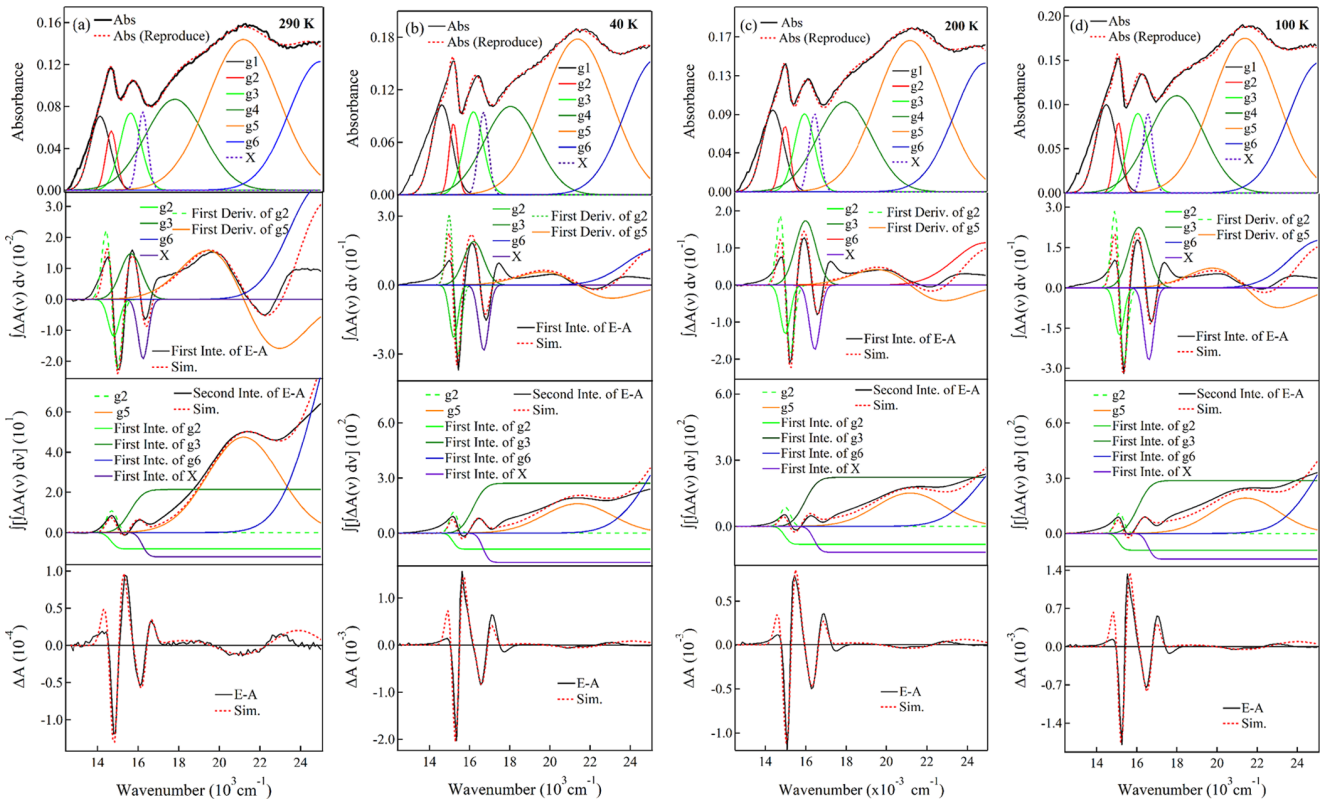


Figure 5. Integral method analysis of the E-A spectra of the MoS_2 thin film obtained at 290 K (a), 40 K (b), 200 K (c), and 100 K (d). From top to bottom, the absorption spectrum along with the Gaussian bands g1−X, the first integral of the E-A spectrum and first derivative of the g2 and g5 spectrum along with g2, g3, g6, and X bands, the second integral of the E-A spectrum and first integral of g2, g3, g6, and X bands along with g2 and g5 bands, and the E-A spectra along with the simulated spectra (bottom).

https://doi.org/10.1021/acs.jpcc.1c06706
J. Phys. Chem. C XXXX, XXX, XXX−XXX

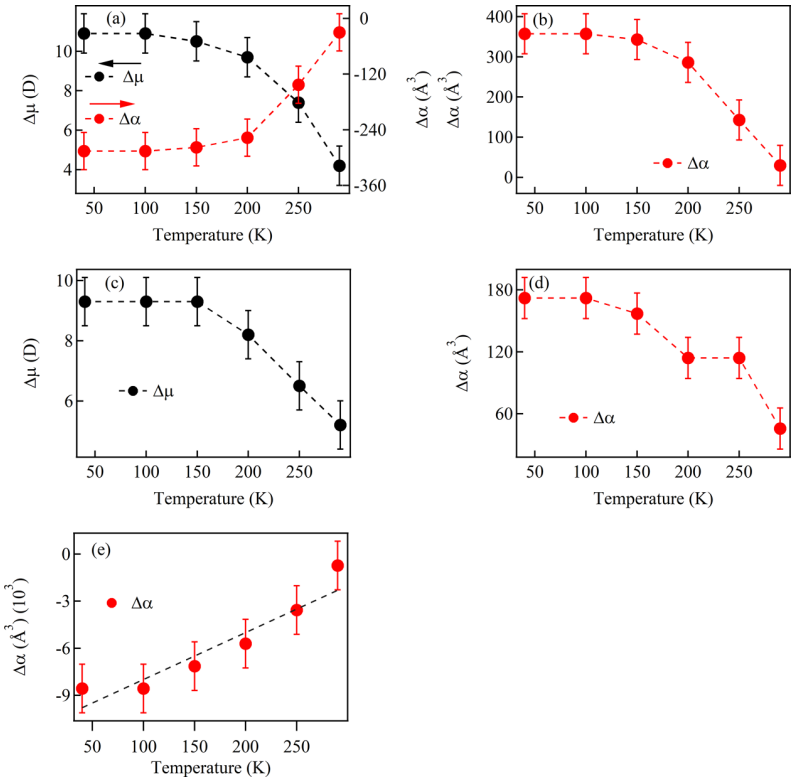


Figure 6. Plots of $|\Delta\mu|$ and $\Delta\bar{\alpha}$ as a function of temperature determined with the integral method analysis of the E-A spectra for the g2 (a), g3 (b), g5 (c), g6 (d), and X (e) bands of the MoS_2 thin film.

derivatives of the absorption intensity $A(\nu)$, as shown below.[17−19]

$$\Delta A = (fF)^2\left[A_\chi A(\nu) + B_\chi \nu \frac{\mathrm{d}}{\mathrm{d}\nu}\left(\frac{A(\nu)}{\nu}\right) + C_\chi \nu \frac{\mathrm{d}^2}{\mathrm{d}\nu^2}\left(\frac{A(\nu)}{\nu}\right)\right] \quad (1)$$

where $F = |\boldsymbol{F}|$ and f denotes the internal field factor. A_χ is a zeroth derivative coefficient and drives the field-induced change in absorption intensity. The coefficient B_χ is responsible for the spectral shift of absorption bands and arises mainly due to $\Delta\alpha$ following absorption. At the same time, C_χ is accountable for a spectral broadening of the absorption bands, which occurs mainly due to $\Delta\mu$. The parameters $\Delta\alpha$ and $\Delta\mu$ can be described in terms of the coefficients B_χ and C_χ as follows

$$B_\chi = \frac{\Delta\bar{\alpha}}{2hc} \text{ and } C_\chi = \frac{|\Delta\mu|^2}{6\,h^2c^2} \quad (2)$$

where $\Delta\bar{\alpha}$ represents the trace of $\Delta\alpha$ (i.e., $\Delta\bar{\alpha} = (1/3)\,\mathrm{Tr}(\Delta\alpha)$). The absorption spectra of MoS_2 NSs represent the overlap of excitonic bands, and the separation of each band is difficult. Therefore, an integral method is considered more advantageous for the E-A spectrum analysis.[26,27] If an E-A spectrum is stated by eq 1, its integral along the wavenumber is given as follows

$$\int \Delta A(\nu)\mathrm{d}\nu \cong (fF)^2\left[A_\chi \int A(\nu)\mathrm{d}\nu + B_\chi A(\nu) + C_\chi \frac{\mathrm{d}A(\nu)}{\mathrm{d}\nu}\right] \quad (3)$$

Further integrating eq 3 over the wavenumber gives

$$\int\left\{\int \Delta A(\nu)\mathrm{d}\nu\right\}\mathrm{d}\nu \cong (fF)^2\left[A_\chi \int\left\{\int A(\nu)\mathrm{d}\nu\right\}\mathrm{d}\nu + B_\chi \int A(\nu)\mathrm{d}\nu + C_\chi A(\nu)\right] \quad (4)$$

The first and second terms on the right-hand side of eq 4 show monotonic increase/decrease as a function ν, depending on the sign of the coefficients A_χ and B_χ. The third term in the right-hand side of eq 4 is equivalent to the absorption spectrum in shape. The absorption spectrum is essential to reproduce with Gaussian bands, i.e., g1−g6 and X, to analyze and reproduce the E-A spectra; the band positions of each Gaussian band are given in Table S1 in the Supporting Information (SI). The peak position, intensity, and band shape, i.e., the bandwidth of these bands, as well as the coefficients A_χ, B_χ, and C_χ, were determined to reproduce the absorption spectrum, E-A spectra, and their first and second integral spectra. These data at different temperatures are summarized in Tables S1 and S2 in the SI. The simulated E-A and integral spectra at 290, 200, 100, and 40 K are shown in Figure 5. The simulation of both E-A and integral spectra at different temperatures in the range of 290−100 K is shown in Figure S1 in the SI. The magnitudes of $|\Delta\mu|$ and $\Delta\bar{\alpha}$ were determined with eq 2 at each temperature in the range of 290−40 K. The results are summarized in Table S3 in the SI. Actually, the X band was not necessary to reproduce the absorption spectra. However, this band was essential to reproduce the E-A spectra, and so the absorption intensity of this band was assumed to be one-twentieth of the

https://doi.org/10.1021/acs.jpcc.1c06706
J. Phys. Chem. C XXXX, XXX, XXX−XXX

g3 band at the peak. The E-A spectra obtained at each temperature could be reproduced with a linear combination of the derivatives of g2, g3, g5, g6, and X bands. The difference in the shape of the E-A spectra between 290 and 40 K, e.g., not only the narrowing of the spectral width but also the presence of a shoulder at a shorter wavelength than the peak at 641 nm at 40 K (see Figure 4a,b), indicates the location of the bands other than two bands (A and B) in this region. Further, it is noted that these bands can be confirmed because not only the E-A spectra but also their first and second integrals are well reproduced by the simulation with an integral method analysis.

With decreasing temperature from 290 to 100 K, the magnitude of $|\Delta\mu|$ and $\Delta\bar{\alpha}$ increases for the g2 band, i.e., 4.2 D and −30 Å^3 at 290 K and 10.9 D and −286 Å^3 at 100 K. For the g3 band, the $\Delta\bar{\alpha}$ value increases with decreasing temperature, i.e., 30 Å^3 at 290 K and 357 Å^3 at 100 K. The second and first derivatives of the absorption spectrum mainly contribute to the E-A spectrum for g5 and g6 bands, respectively. $|\Delta\mu|$ for g5 and $\Delta\bar{\alpha}$ for g6 increase with decreasing temperature, i.e., 5.2 D at 290 K and 9.3 D at 100 K for g5 and 46 Å^3 at 290 K and 172 Å^3 at 100 K for g6. The first derivative of the X band has a significant contribution to the E-A spectrum, indicating that the magnitude of $\Delta\bar{\alpha}$ increased with decreasing temperature, i.e., −728 Å^3 at 290 K and −8568 Å^3 at 100 K. g1 and g4 bands do not contribute to the E-A spectrum at any temperature. For g2, g3, g6, and X bands, the plots of $\Delta\bar{\alpha}$ versus temperature show a similar relation, i.e., the magnitude of $\Delta\bar{\alpha}$ increased monotonically with decreasing temperature. The magnitude of $|\Delta\mu|$ also increases with decreasing temperature, as observed for g2 and g5 bands. The variations in magnitudes of $|\Delta\mu|$ and $\Delta\bar{\alpha}$ as a function of temperature are plotted for g2, g3, g5, g6, and X bands, as shown in Figure 6. Both $|\Delta\mu|$ and $\Delta\bar{\alpha}$ values are found to have a maximum at 100 K and nearly constant thereafter. Thus, both $|\Delta\mu|$ and $\Delta\bar{\alpha}$, which are directly related to the charge separation and charge distribution following photoexcitation, respectively, are very sensitive to the temperature. The large value of $|\Delta\mu|$ implies a charge transfer from chalcogen (S) to Mo following absorption, which substantially enhances with decreasing temperature, as observed for g2 and g5 bands. Actually speaking, the weak negative peak observed at low temperatures in the region of 17 000−18 500 cm^{-1} could not be reproduced (see Figure 5 at 40 K and Figure S1 in the SI). This peak was not observed at a high temperature of 290 K (see Figure 4a). As shown in Figure 4b, the spectral shape in this region depends on the applied field strength, and the zero-crossing point at around 572 nm as well as the minimum position located between 560 and 570 nm shows a blue shift with increasing applied field strength. It is also noted that a positive signal appeared in the shorter wavelength region compared to the zero crossing at around 545 nm, though the shape of the E-A spectra caused by the Stark effect given by eq 1 is considered to be independent of the applied field strength.

The observed absorption spectra were reproduced with a linear combination of six Gaussian bands, but only g2, g3 g5, g6, and X bands are required to reproduce the E-A spectra. The g1 and g4 bands have a negligible contribution to the E-A spectrum, whereas the band X has a substantial contribution. The g1, g2, g3, and X bands show a blue shift with decreasing temperature, but the separation between g1 and g2 (about 26 ± 3 nm) and between g3 and X (22 ± 3 nm) is nearly independent of the temperature, as shown in Table S1 in the SI. As the origin of the spectral broadening of each band having a Gaussian profile, the nonradiative process such as exciton dissociation can be considered. As the dissociation process following photoexcitation becomes faster, the spectral width becomes larger.

Vella et al.[30] reported that the exciton band "A" observed at ∼660 nm is an overlapping of two electronic resonance bands, namely, (A^0) and $(\bar{A})$, based on the analysis of the E-A spectra with the derivative method. Then, g1, g2, g3, and X bands confirmed in the present study may be attributed to the splitting of A and B excitonic bands and assigned to the $(\bar{A})$, (A^0) and $(\bar{B})$, (B^0) resonances, respectively, which behave differently with the applied field. The g1 band assigned to the $(\bar{A})$ feature does not contribute to the overall E-A spectrum. However, the relatively narrow band (g2) corresponding to the exciton contributes significantly to the E-A spectrum. Likewise, g3 and X bands corresponding to the $(\bar{B})$ and (B^0) features substantially contribute to the E-A spectrum. It is stressed that the exciton X band, which is not clear in the absorption spectra, contributes largely to the E-A spectra. In monolayer MoS_2 NS, the exciton band B of the absorption spectrum consists of two overlapping electronic resonances, i.e., (B^0) and $(\bar{B})$, where (B^0) and $(\bar{B})$ may correspond to the transition from the neutral ground state to a neutral exciton state and from a charged ground state to a charged exciton state (trion), respectively.[28,30] Thus, it has been concluded that in multilayer MoS_2 NSs, both A and B excitonic bands consist of the two resonance bands, which depend on the applied electric field. For example, the $|\Delta\mu|$ and $\Delta\bar{\alpha}$ values of corresponding transitions increase substantially with decreasing temperature. The $|\Delta\mu|$ value increased about 2 times, whereas $\Delta\bar{\alpha}$ increased more than 10 times. These values do not further increase upon decreasing the temperature below 100 K. It can be mentioned that the amount of charge transfer following photoexcitation increases significantly upon decreasing the temperature.

CONCLUSIONS

MoS_2 NSs were synthesized with a hydrothermal method. The crystallinity and layered structure of the MoS_2 NSs were confirmed with HRTEM images and temperature-dependent XRD patterns. The layers of MoS_2 NSs are well stacked with an interlayer distance of 0.66 nm. The absorption and E-A spectra of MoS_2 NSs as a function of temperature (290−40 K) and electric field strength (0.1−0.5 MV cm^{-1}) have been measured under vacuum conditions, and E-A spectra were analyzed with an integral method by assuming that the E-A signal arises from the quadratic Stark effect. Both A and B excitonic bands split into the resonance states. The magnitudes of $|\Delta\mu|$ and $\Delta\bar{\alpha}$ for excitonic resonances increased with decreasing temperature and became constant at around 100 K. The X band exhibits massive changes in $\Delta\bar{\alpha}$, which depends on temperature, possibly as a result of temperature-dependent lattice distortion. The study of the electric field-induced change in absorption spectra is important since the application of electric field is essential in optoelectronic device operation.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcc.1c06706.

Gaussian band positions used to reproduce the absorption spectra; values of coefficients of A_χ, B_χ, and C_χ obtained from the simulation of E-A spectra; values

https://doi.org/10.1021/acs.jpcc.1c06706
J. Phys. Chem. C XXXX, XXX, XXX−XXX

of $|\Delta\mu|$ and $\Delta\bar{\alpha}$ estimated from the simulation of the E-A spectra; and absorption and E-A spectra of the MoS_2 thin film at 250 and 150 K and their simulation with an integral method analysis (PDF)

AUTHOR INFORMATION

Corresponding Authors

Mohan Singh Mehata − *Laser-Spectroscopy Laboratory, Department of Applied Physics, Delhi Technological University, Delhi 110042, India;* orcid.org/0000-0003-4564-5243; Email: msmehata@gmail.com

Nobuhiro Ohta − *Department of Applied Chemistry and Institute of Molecular Science and Center for Emergent Functional Matter Science, National Yang Ming Chiao Tung University, Hsinchu 300093, Taiwan;* orcid.org/0000-0003-4255-6448; Email: nohta@nctu.edu.tw

Authors

Mrityunjay Kumar Singh − *Laser-Spectroscopy Laboratory, Department of Applied Physics, Delhi Technological University, Delhi 110042, India*

Kamlesh Awasthi − *Department of Applied Chemistry and Institute of Molecular Science and Center for Emergent Functional Matter Science, National Yang Ming Chiao Tung University, Hsinchu 300093, Taiwan;* orcid.org/0000-0001-7852-059X

Prateek Sharma − *Laser-Spectroscopy Laboratory, Department of Applied Physics, Delhi Technological University, Delhi 110042, India*

Shailesh Rana − *Department of Applied Chemistry and Institute of Molecular Science, National Yang Ming Chiao Tung University, Hsinchu 300093, Taiwan*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcc.1c06706

Author Contributions

‖M.S.M. and M.K.S. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

M.S.M. and M.K.S. thank National Yang Ming Chiao Tung University (NYCU), Taiwan, for supporting their stay at NYCU. Taiwan Ministry of Science and Technology (MOST 108-2113-M-009-002 and MOST 108-2113-M-009-001) and Center for Emergent Functional Matter Science of National Yang Ming Chiao Tung University from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by Taiwan Ministry of Education (MOE) supported this work.

REFERENCES

(1) Zhou, W.; Zou, X.; Najmaei, S.; Liu, Z.; Shi, Y.; Kong, J.; Lou, J.; Ajayan, P. M.; Yakobson, B. I.; Idrobo, J. C. Intrinsic Structural Defects in Monolayer Molybdenum Disulfide. *Nano Lett.* **2013**, *13*, 2615−2622.

(2) Jaiswal, H. N.; Liu, M.; Shahi, S.; Yao, F.; Zhao, Q.; Xu, X.; Li, H. Localized Surface Plasmon Resonance on Two-Dimensional $HfSe_2$ and $ZrSe_2$. *Semicond. Sci. Technol.* **2018**, *33*, No. 124014.

(3) Sharma, P.; Mehata, M. S. Colloidal MoS_2 Quantum Dots Based Optical Sensor for Detection of 2,4,6-TNP Explosive in an Aqueous Medium. *Opt. Mater.* **2020**, *100*, No. 109646.

(4) Sharma, P.; Mehata, M. S. Rapid Sensing of Lead Metal Ions in an Aqueous Medium by MoS_2 quantum Dots Fluorescence Turn-Off. *Mater. Res. Bull.* **2020**, *131*, No. 110978.

(5) Pham, T.; Li, G.; Bekyarova, E.; Itkis, M. E.; Mulchandani, A. MoS_2-Based Optoelectronic Gas Sensor with Sub-Parts-per-Billion Limit of NO_2 Gas Detection. *ACS Nano* **2019**, *13*, 3196−3205.

(6) Kim, Y.; Bark, H.; Kang, B.; Lee, C. Wafer-Scale Substitutional Doping Of Monolayer MoS_2 Films For High-Performance Optoelectronic Devices. *ACS Appl. Mater. Interfaces* **2019**, *11*, 12613−12621.

(7) Parzinger, E.; Miller, B.; Blaschke, B.; Garrido, J. A.; Ager, J. W.; Holleitner, A.; Wurstbauer, U. Photocatalytic Stability Of Single- And Few-Layer MoS_2. *ACS Nano* **2015**, *9*, 11302−11309.

(8) Singh, M. K.; Mehata, M. S. Enhanced Photoinduced Catalytic Activity of Transition Metal Ions Incorporated TiO_2 Nanoparticles for Degradation of Organic Dye: Absorption and Photoluminescence Spectroscopy. *Opt. Mater.* **2020**, *109*, No. 110309.

(9) Singh, M. K.; Mehata, M. S. Phase-Dependent Optical and Photocatalytic Performance of Synthesized Titanium Dioxide (TiO_2) Nanoparticles. *Optik* **2019**, *193*, No. 163011.

(10) Lee, H. S.; Min, S.-W.; Chang, Y.-G.; Park, M. K.; Nam, T.; Kim, H.; Kim, J. H.; Ryu, S.; Im, S. MoS_2 Nanosheet Phototransistors with Thickness-Modulated Optical energy Gap. *Nano Lett.* **2012**, *12*, 3695−3700.

(11) Lopez-Sanchez, O.; Lembke, D.; Kayci, M.; Radenovic, A.; Kis, A. Ultrasensitive Photodetectors Based on Monolayer MoS_2. *Nat. Technol.* **2013**, *8*, 497−501.

(12) Tsai, M.-L.; Su, S.-H.; Chang, J.-K.; Tsai, D.-S.; Chen, C.-H.; Wu, C.-I.; Li, L.-J.; Chen, L-J.; He, J.-H. Monolayer MoS_2 Heterojunction Solar Cells. *ACS Nano* **2014**, *8*, 8317−8322.

(13) Frindt, R. F.; Yoffe, A. D. Physical Properties of Layer Structures: Optical Properties and Photoconductivity of Thin Crystals of Molybdenum Disulphide. *Proc. R. Soc. London, Ser. A* **1962**, *273*, 69−83.

(14) Evans, B. L.; Young, P. A. Optical Absorption and Dispersion in Molybdenum Disulphide. *Proc. R. Soc. London, Ser. A* **1965**, *284*, 402−422.

(15) Coehoorn, R.; Haas, C.; Dijkstra, J.; Flipse, C. J. F.; de Groot, R. A.; Wold, A. Electronic Structure of $MoSe_2$, MoS_2, and WSe_2. I, Band-Structure Calculations and Photoelectron Spectroscopy. *Phys. Rev. B* **1987**, *35*, 6195−6202.

(16) Liu, G.; Ma, H.; Teixeira, I.; Sun, Z.; Xia, Q.; Hong, X.; Tsang, S. C. E. Hydrazine-Assisted Liquid Exfoliation of MoS_2 for Catalytic Hydrodeoxygenation of 4-Methylphenol. *Chem. Eur. J.* **2016**, *22*, 2910−2914.

(17) Liptay, W. Dipole Moments and Polariabilities of Molecules in Excited Electronic States. In *Excited States*; Lim, E. C., Ed.; Elsevier: New York, USA, 1974; Vol. *1*, pp 129−229.

(18) Bublitz, G. U.; Boxer, S. G. Stark Spectroscopy: Applications in Chemistry, Biology, and Materials Science. *Annu. Rev. Phys. Chem.* **1997**, *48*, 213−242.

(19) Jalviste, E.; Ohta, N. Theoretical Foundation of Electroabsorption Spectroscopy: Self-Contained Derivation of the Basic Equations with the Direction Cosine Method and the Euler Angle Method. *J. Photochem. Photobiol., C* **2007**, *8*, 30−46.

(20) Umeuchi, S.; Nishimura, Y.; Yamazaki, I.; Murakami, H.; Yamashita, M.; Ohta, N. Electric Field Effects on Absorption and Fluorescence Spectra of Pyrene Doped in a PMMA Polymer Film. *Thin Solid Films* **1997**, *311*, 239−245.

(21) Locknar, S. A.; Peteanu, L. A. Investigation of the Relationship between Dipolar Properties and Cis-Trans Configuration in Retinal Polyenes: A Comparative Study Using Stark Spectroscopy and Semiempirical Calculations. *J. Phys. Chem. B* **1998**, *102*, 4240−4246.

(22) Nakabayashi, T.; Wahadoszamen, M.; Ohta, N. External Electric Field Effects on the Photoexcitation Dynamics of Diphenylpolyenes. *J. Am. Chem. Soc.* **2005**, *127*, 7041−7052.

(23) Mehata, M. S.; Hsu, C. S.; Lee, Y. P.; Ohta, N. Electroabsorption and Electrophotoluminescence of Poly(2,3-Diphenyl-5-Hexyl-p-Phenylene Vinylene). *J. Phys. Chem. C* **2012**, *116*, 14789−14795.

(24) Awasthi, K.; Du, K. B.; Wang, C. Y.; Tsai, C. L.; Hamada, M.; Narra, S.; Diau, E. W. G.; Ohta, N. Electroabsorption Studies of Multicolored Lead Halide Perovskite Nanocrystalline Solid Films. *ACS Photonics* **2018**, *5*, 2408−2417.
(25) Rana, S.; Awasthi, K.; Bhosale, S. S.; Diau, E. W. G.; Ohta, N. Temperature-Dependent Electroabsorption and Electrophotoluminescence and Exciton Binding Energy in $MAPbBr_3$ Perovskite Quantum Dots. *J. Phys. Chem. C* **2019**, *123*, 19927−19937.
(26) Awasthi, K.; Iimori, T.; Ohta, N. Integral Method Analysis of Electroabsorption Spectra and Its Application to Quantum Dots of PbSe. *J. Phys. Chem. C* **2014**, *118*, 18170−18176.
(27) Kattoor, V.; Awasthi, K.; Jokar, E.; Diau, E. W. G.; Ohta, N. Integral Method Analysis of Electroabsorption Spectra and Electrophotoluminescence Study of $(C_4H_9NH_3)_2PbI_4$ Organic-Inorganic Quantum Well. *J. Phys. Chem. C* **2018**, *122*, 26623−26634.
(28) Mak, K. F.; He, K.; Lee, C.; Lee, G. H.; Hone, J.; Heinz, T. F.; Shan, J. Observation of Tightly Bound Trions in Monolayer MoS_2. *Nat. Mater.* **2013**, *12*, 207−211.
(29) Bordas, J.; Davis, E. A. Electromodulation Spectroscopy of Excitons: Molybdenum Disulphide. *Phys. Status Solidi B* **1973**, *60*, 505−510.
(30) Vella, D.; Ovchinnikov, D.; Martino, N.; Vega-Mayoral, V.; Dumcenco, D.; Kung, Y.-C.; Antognazza, M.-R.; Kis, A.; Lanzani, G.; Mihailovic, D.; et al. Unconventional Electroabsorption in Monolayer MoS_2. *2D Mater.* **2017**, *4*, No. 021005.
(31) Shi, J.; Baldini, E.; Latini, S.; Sato, S. A.; Zhang, Y.; Pein, B. C.; Shen, P.-C.; Kong, J.; Rubio, A.; Gedik, N.; et al. Room Temperature Terahertz Electroabsorption Modulation by Excitons in Monolayer Transition Metal Dichalcogenides. *Nano Lett.* **2020**, *20*, 5214−5220.
(32) Chaudhary, N.; Khanuja, M.; Abid; Islam, S. S. Hydrothermal Synthesis of MoS_2 Nanosheets for Multiple Wavelength Optical Sensing Applications. *Sens. Actuators, A* **2018**, *277*, 190−198.
(33) Beal, A. R.; Knights, J. C.; Liang, W. Y. Transmission Spectra of Some Transition Metal Dichalcogenides. I. Group IVA: Octahedral Coordination. *J. Phys. C: Solid State Phys.* **1972**, *5*, 3531−3539.
(34) Murray, R. B.; Yoffe, A. D. The Band Structures of Some Transition Metal Dichalcogenides: Band Structures of the Titanium Dichalcogenides. *J. Phys. C: Solid State Phys.* **1972**, *5*, 3038−3046.
(35) Wang, K.; Wang, J.; Fan, J.; Lotya, M.; O'Neill, A.; Fox, D.; Feng, Y.; Zhang, X.; Jiang, B.; Zhao, Q.; et al. Ultrafast Saturable Absorption of Two-Dimensional MoS_2 Nanosheets. *ACS Nano* **2013**, *7*, 9260−9267.
(36) Ellis, J. K.; Lucero, M. J.; Scuseria, G. E. The Indirect to Direct Band Gap Transition in Multilayered MoS_2 as Predicted by Screened Hybrid Density Functional Theory. *Appl. Phys. Lett.* **2011**, *99*, No. 261908.
(37) Yalon, E.; Aslan, Ö. B.; Smithe, K. K. H.; McClellan, C. J.; Suryavanshi, S. V.; Xiong, F.; Sood, A.; Neumann, C. M.; Xu, X.; Goodson, K. E.; et al. Temperature-Dependent Thermal Boundary Conductance of Monolayer MoS_2 by Raman Thermometry. *ACS Appl. Mater. Interfaces* **2017**, *9*, 43013−43020.

https://doi.org/10.1021/acs.jpcc.1c06706
J. Phys. Chem. C XXXX, XXX, XXX−XXX

Semiconductor Science and Technology

ACCEPTED MANUSCRIPT

Theoretical study of highly efficient CH3NH3SnI3 based Perovskite solar cell with CuInS2 quantum dot

To cite this article before publication: Gagandeep Kundu *et al* 2021 *Semicond. Sci. Technol.* in press https://doi.org/10.1088/1361-6641/ac4325

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 22/12/2021 at 06:34

Theoretical Study of Highly Efficient $CH_3NH_3SnI_3$ based Perovskite Solar Cell with $CuInS_2$ Quantum Dot

Gagandeep[1], Mukhtiyar Singh[2], Ramesh Kumar[1, a)] ,Vinamrita Singh[3] and Sunita Srivastava[1,4]
[1]Department of Physics, Guru Jambheshwar University of Science & Technology, Hisar, India
[2]Department of Applied Physics, Delhi Technological University, Delhi, India
[3]Department of Applied Sciences and Humanities, Netaji Subhas University of Technology – East Campus (Formerly AIACT&R), Delhi, India
[4]Department of Physics, Panjab University Chandigarh, India

[a)]Corresponding author: ramesh85@gjust.org

Abstract:

Simulation studies have been carried out for the n-i-p perovskite solar cell (PSC) structure i.e. ITO/SnO_2/$CH_3NH_3PbI_3$/$CuInS_2$/Au. We have considered this cell as our primary structure and is simulated using Solar Cell Capacitance Simulator (SCAPS-1D) software. Here, the $CuInS_2$ quantum dot acts as an inorganic hole transporting layer. Further, the use of the $CuInS_2$ quantum dot in PSCs has been explored by simulating twenty different cell structures. These perovskite solar cells are based on recently used absorber layers, i.e., $MASnI_3$, $FAPbI_3$, and $(FAPbI_3)_{0.97}(MAPbBr_{1.5}Cl_{1.5})_{0.03}$, and electron transporting layers, i.e., SnO_2, TiO_2, ZnO, C_{60}, and IGZO. The performance of all structures has been optimized by varying the thickness of the absorber layers and ETLs. The cell structure, ITO/SnO_2/$CH_3NH_3SnI_3$/$CuInS_2$/Au, has been found to exhibit the highest power conversion efficiency of 21.79 % as compared to other cells. Investigations have also been carried out to analyze the effect of defect density in the absorber layer and the interface of the cell structure. In addition, the cell performance has been ascertained by examining the impact of operating temperature, metal contact work function and that of resistance in series as well as in parallel. The simulation results of our primary cell structure are found to be in good agreement with the recent experimental study.

Keywords: Perovskite Solar Cell, Quantum Dot, Hole Transporting Layer, SCAPS, Efficiency, Temperature, CIS QD.

Introduction: The PSC based on organic-inorganic hybrid lead (Pb) halide perovskites show tremendous potential for developing efficient and cost-effective solar cell technology. The perovskites based on $CH_3NH_3PbI_3$ (methyl ammonium lead iodide or MAPI) material have remarkable optoelectronic properties such as high value of absorption coefficient (10^3-10^4 cm^{-1}), long carrier diffusion length (0.1-1.9 μm), high value of carrier mobilities (2-66 cm^2 V^{-1} s^{-1}), low exciton binding energy ($\leq$ 5 meV), and tunable band gaps (1.55-2.5 eV) [1-2].The power conversion efficiency (PCE) of the PSCs based on MAPI has been observed to increase at a remarkable rate, i.e., 3.8 % in 2009 [3] to 25.2 % in 2020 [4].

Accepted Manuscript

PSCs generally have a device structure comprised of a perovskite absorber layer sandwiched between the electron transporting layer (ETL) and the hole transporting layer (HTL). The most commonly used absorber layer is based on MAPI. However, its toxicity and instability may lower its potential use for commercialization in PSCs [5]. Hence, there is a need to investigate alternate absorber layers that enhance stability and reduce PSC toxicity while exhibiting a comparable performance.

The formamidinum lead iodide perovskite ($HC(NH_2)_2PbI_3$ or $FAPbI_3$) has been used as an absorber layer in PSCs [6]. It received great attention compared to MAPI due to its optimum bandgap (1.3-1.48 eV). This bandgap value lies closer to the infra-red single-junction range [7].The $FAPbI_3$ based PSC was recently studied with a PCE reported value of 21.4 % [6].

The use of organic/inorganic cation and/or mixed anion are being investigated to enhance PSCs stability. Chen et al. [8] investigated the PSC based on $(FAPbI_3)_{0.97}(MAPbBr_{1.5}Cl_{1.5})_{0.03}$ or (FAPI)(MAPBC) as an absorber and obtained a PCE of 21.1 %.

Although the above-mentioned absorber layers enhanced the PSCs stability, the issue of toxicity still remains due to the presence of Pb. The tin-based perovskite ($CH_3NH_3SnI_3$ or MASI) is another Pb-free material used in PSCs structures [9-11]. However, the performance of Pb-free PSCs is still far behind as compared to MAPI based solar cells. Theoretical and experimental studies of MASI based solar cells have efficiency in the range of 6 % to 18.71 % [12-13].

In addition to the choice of the different kinds of absorber layer, the ETLs and HTLs have been found to significantly affect the stability and performance of PSCs. The most commonly used organic and inorganic ETL materials are organic fullerene derivative [6, 6] - phenyl-C_{61}-butyric acid methyl ester (PCBM), tin dioxide (SnO_2) [14], zinc oxide (ZnO) [15], indium gallium zinc oxide (IGZO) [16], C_{60} [17], and titanium oxide (TiO_2) [18]. The most used organic HTLs in PSCs are spiro-OMeTAD (2, 2', 7, 7'-tetrakis [N, N- di (4-methoxyphenyl) amino]-9, 9'-spirobifluorene) [19], nanofibular polyaniline (PANI) [20], and polytriarylamine (PTAA) [21]. In the case of inorganic HTLs, CuI, NiO, CuSCN and CuS materials have been used [22-23]. Cu-based chalcopyrite semiconductors such as $CuInSe_2$, $CuInS_2$, Cu_2ZnSnS_4, $Cu_2ZnSnSe_4$, and $Cu(In,Ga)(S,Se)_2$, have also been realized in PSCs as inorganic HTLs [24-27].

The $CuInS_2$, a ternary semiconductor as compared to other Cu based chalcopyrite semiconductors, has some outstanding properties such as low toxicity, appropriate bandgap for absorption, and large extinction coefficient [28]. A $CuInS_2$ quantum dot (CIS QD) was also implemented in PSCs as inorganic HTLs with a reported PCE of 8.38 % was obtained [25].

Recently, Liua et al. [29] fabricated the CIS QD using a modified hot injection method. The fabricated CIS QD was used as a HTL with SnO_2 ETL and MAPI absorber in the PSC structure, i.e., ITO/SnO_2/$CH_3NH_3PbI_3$/$CuInS_2$/Au. They obtained a PCE of 18.8 % and demonstrated the improved stability of the cell as compared to Spiro-OMeTAD HTL. The organic ETLs and HTLs have their limitations, such as complicated fabrication steps and hygroscopic nature that causes less stability as compared to their inorganic counterparts.

With this perspective, we have taken the experimentally studied PSC structure [29] as our primary structure and simulated it using SCAPS-1D software [30]. The obtained data for the primary

Accepted Manuscript

structure has been compared with the experimental results. Consequently, a search has been carried out to find the most efficient Pb-free i.e less toxic PSC structure.

To the best of our awareness of the literature, any detailed report on the simulation of CIS QD as HTL in PSCs is not available. Therefore, we simulated twenty PSC structures with recently used absorbers, i.e., MASI [13], $FAPbI_3$ [6], and (FAPI)(MAPBC) [10] in the primary structure and SnO_2, TiO_2, ZnO, C_{60}, and IGZO as ETL with CIS QD based HTL. All the simulated structures have been optimized by varying the absorbers and ETLs thickness, keeping other parameters fixed. The simulation results provided the best-optimized PSC device with the highest performance, i.e., ITO/SnO_2/$CH_3NH_3SnI_3$/$CuInS_2$/Au. The aim of our study is to establish a lead-free PSC structure with good performance. Furthermore, we investigated the effect of defect density in MASI absorber and at ETL/absorber/HTL interfaces.

The highest efficiency cell structure is further optimized by studying the effect of various metals as back contact electrodes, and the effect of series and parallel resistances (R_s and R_p). The effect of the operating temperature has also been discussed. This work provides the basic guideline for the researchers to select the most efficient lead-free PSC.

Device structure and methodology: The primary cell structure of ITO/SnO_2/$CH_3NH_3PbI_3$/$CuInS_2$/Au was taken from the experimental study [29].The excitons are produced when the light falls on the MAPI layer. Electrons move towards the SnO_2 layer and the holes move towards the CIS QD. The charge carriers are collected by the respective front contact (ITO) and the metal back contact (Au).

To explore the use of CIS QD as HTL in PSCs, the MAPI absorber has been replaced by the following different absorbers: MASI, $FAPbI_3$, and (FAPI)(MAPBC). The following ETLs have been used for the simulations: SnO_2, TiO_2, ZnO, C_{60} and IGZO. Fig.1 depicts the energy band diagram for the different cell structures. All the PSC structures were modelled under the solar irradiance of AM 1.5 G with an illumination intensity of 1000W/m^2 at 300 K operating temperature. The Working of the software is based on three semiconductor equations: Poisson's equation, carrier continuity equation, and drift-diffusion equations.

The Poisson's equation is written as

$$\frac{\partial^2\Psi}{\partial^2 x} = -\frac{\partial E}{\partial x} = -\frac{\rho}{\varepsilon_s} = -\frac{q}{\varepsilon_s}\left[p - n + N_D^+(x) - N_A^-(x) \pm N_{def}(x)\right] \quad 1.$$

here Ψ is the electrostatic potential, E represents the electric field of the p-n junction, q is the charge, ε_s is the static relative permittivity of the medium, p(n) is the hole (electron) density, $N_A^-(N_D^+)$ is the density of ionized acceptors (donors), and N_{def} is the possible defect density of acceptor and donor [31]. Hence, this Poisson equation gives the relation between electric field and space charge density (ρ).

The hole and electron continuity equations in steady state are given as:

$$-\frac{\partial j_p}{\partial x} + G - U_p(n,p) = 0 \quad 2.$$

$$\frac{\partial j_n}{\partial x} + G - U_n(n,p) = 0 \quad 3.$$

Accepted Manuscript

Here, j_p and j_n are hole and electron densities; G is the electron-hole recombination rate, and $U_{n,p}$ is the net recombination rates.

The current density of hole and electron is given by:

$$j_p = qp\mu_p E - qD_p \frac{\partial p}{\partial x} \quad (4.)$$

$$j_n = qp\mu_n E - qD_n \frac{\partial n}{\partial x} \quad (5.)$$

$\mu_{p(n)}$ is the hole (electron) mobility, q is the charge, and $D_{p(n)}$ is the diffusion coefficient of the hole (electron).

The cell parameters employed in this simulation have been taken from recently reported theoretical and experimental studies [18, 29, 32-42]. The corresponding data has been provided in tables 1 and 2. Here, acceptor and donor densities are represented by N_A and N_D, dk is the dielectric constant, E_g is the bandgap energy, χ is the electron affinity, μ_p and μ_n are the mobility of holes and electrons, and N_t denotes the defect density. Initially, the behavior of the perovskite absorber layer is almost intrinsic.

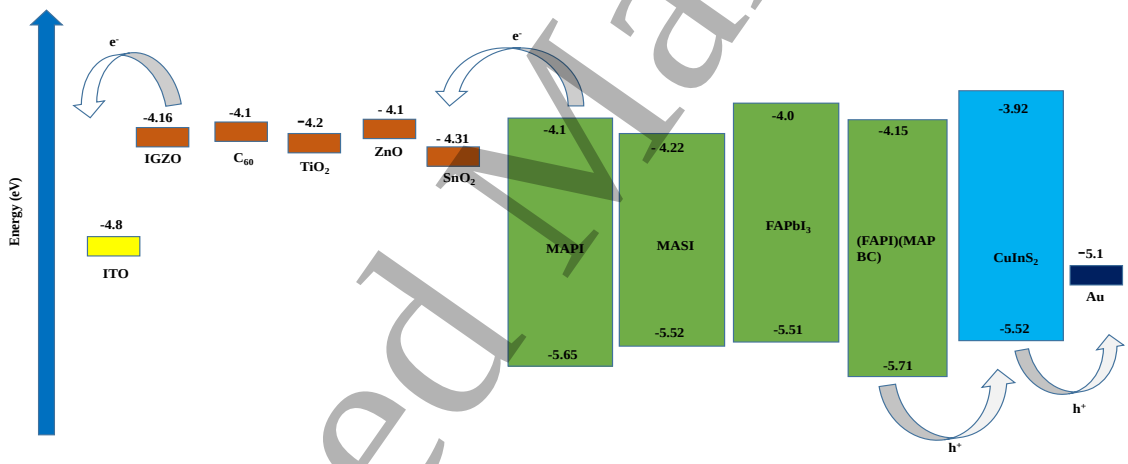


Fig. 1. Energy band diagram of different materials used in the simulation of perovskite solar cells with diverse ETLs.

The defect density (N_t) for the perovskite layer was set at 4.8×10^{14} cm^{-3}, giving a carrier diffusion length equal to 1 µm for holes and electrons. This value is quite similar to the values given in literature [43]. Some of these parameters, not mentioned in table 1, are identical for all the layers; the thermal velocity of holes and electrons is 1×10^7 cm/s. The defect type is kept neutral; the defect energy level is centred at the E_g with a characteristics energy of 1.0 eV and an energy distribution of Gaussian.

Accepted Manuscript

Table 1. Modeling parameters used for the different materials in PSCs

Parameters	ITO	SnO_2	MAPI	MASI	$FAPbI_3$	(FAPI)(MAPBC)	CIS QD
Thickness (nm)	60	variable	variable	variable	variable	variable	150
ε_r	10	9	6.5	6.5	6.6	6.5	9.76
E_g(eV)	3.6	3.79	1.55	1.3	1.51	1.56	1.6
χ(eV)	4.1	4.0	3.9	4.17	4.0	3.9	4.1
N_c (cm^{-3})	2.2×10^{18}	2.2×10^{18}	2.5×10^{19}	1×10^{18}	1.2×10^{19}	2.2×10^{18}	2×10^{18}
N_v (cm^{-3})	1.8×10^{19}	1.8×10^{19}	2.5×10^{19}	1×10^{19}	2.9×10^{18}	1.9×10^{19}	2×10^{19}
μ_e ($cm^2\ V^{-1}\ s^{-1}$)	50	240	2	1.6	2.7	2	100
μ_h($cm^2\ V^{-1}\ s^{-1}$)	75	15	2	1.6	1.8	2	25
v_e(cm/s)	1×10^7	1×10^7	1×10^7	1×10^7	1×10^7	1×10^7	1×10^7
v_h(cm/s)	1×10^7	1×10^7	1×10^7	1×10^7	1×10^7	1×10^7	1×10^7
N_A(cm^{-3})	0	0	5.21×10^9	3.2×10^{15}	1.3×10^{16}	1×10^{13}	0
N_D(cm^{-3})	1×10^{20}	1×10^{15}	5.21×10^9	3.2×10^{15}	1.3×10^{16}	1×10^{13}	2×10^{16}
N_t(cm^{-3})	1×10^{19}	1×10^{15}	4.8×10^{14}	3.9×10^{14}	5.4×10^{14}	4.8×10^{14}	1×10^{14}

Table 2. Material parameters of proposed ETLs

Parameter	TiO_2	ZnO	C_{60}	IGZO
Thickness (nm)	50	50	50	50
ε_r	9	9	4.2	10
E_g(eV)	3.2	3.3	1.7	3.05
χ(eV)	3.9	4.1	3.9	4.16
N_c (cm^{-3})	1×10^{21}	4×10^{18}	8×10^{19}	5×10^{18}
N_v (cm^{-3})	2×10^{20}	1×10^{19}	8×10^{19}	5×10^{18}
μ_e ($cm^2\ V^{-1}\ s^{-1}$)	20	100	8×10^{-2}	15
μ_h($cm^2\ V^{-1}\ s^{-1}$)	10	25	3.5×10^{-3}	0.1
v_e(cm/s)	1×10^7	1×10^7	1×10^7	1×10^7
v_h(cm/s)	1×10^7	1×10^7	1×10^7	1×10^7
N_A(cm^{-3})	0	1×10^5	0	0
N_D(cm^{-3})	1×10^{16}	1×10^{16}	2.6×10^{18}	1×10^{16}
N_t(cm^{-3})	1×10^{15}	2×10^{17}	1×10^{14}	1×10^{15}

3. Results and discussion

3.1 Device optimization: The n-i-p configuration, ITO/SnO_2/$CH_3NH_3PbI_3$/$CuInS_2$/Au, has been analyzed with a layer thickness of 50 nm, 100 nm, and 150 nm for ETL, absorber, and HTL, respectively. The thickness of the ETL layer was varied from 50 nm to 400 nm. The results for the PSC performance with a change in ETL thickness have been shown in Fig. 2. From the results of Fig. 2, it can be clearly noticed that the J_{sc} and PCE decrease as we increase the thickness of the SnO_2 layer. The cell's efficiency has been found to be optimum at 50 nm thickness of SnO_2; which

Accepted Manuscript

decreases thereafter. This is due to an increase in charge recombination rate with the increase in the thickness of the ETL layer as shown in Fig. 3.

After optimizing the ETL thickness, the absorber layer thickness is varied from 100 nm to 2000 nm, and the corresponding results of the solar cell parameters have been plotted in Fig. 4. The results shown in Fig. 4 clearly illustrate that as the absorber layer thickness is increased, a significantly improved performance of the cell is attained. This could be explained on the basis of absorption, i.e., a thicker layer absorbs more photons and creates more electron-hole pairs. The value of J_{sc}, V_{oc}, FF, and PCE increases from 13.85 mA/cm^2 to 24.67 mA/cm^2, 811.10 mV to 858.30 mV, 77.60 % to 81.05 %, and 8.72 % to 17.16 %, respectively, with the absorber thickness increasing from 100 nm to 1000 nm. At 1000 nm absorber thickness, J_{sc} and PCE almost saturate to form the plateau, and ultimately a decrement in the performance is obtained beyond the value of the thickness of 1500 nm. After a certain limit, the recombination rate increases, which suppresses the charge carriers collected at the respective electrode. Keeping in mind the small variation in the cell parameters after 1000 nm and experimental diffusion length, absorber thickness was kept at 1000 nm. There are various studies reported in the literature which give the optimized value of the thickness of more than 500 nm [44, 45].

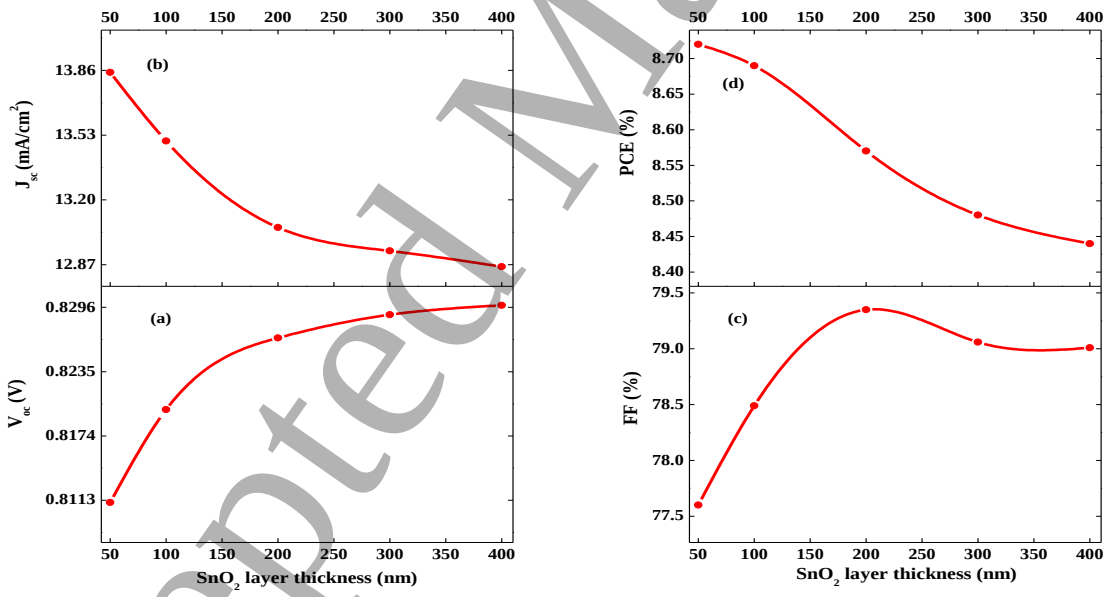


Fig.2. Change in (a) V_{oc}, (b) J_{sc}, (c) FF, and (d) PCE with the ETL thickness of Pb-based PSCs. Thickness of the perovskite absorber layer and HTL are 100 nm and 150 nm, respectively.

Accepted Manuscript

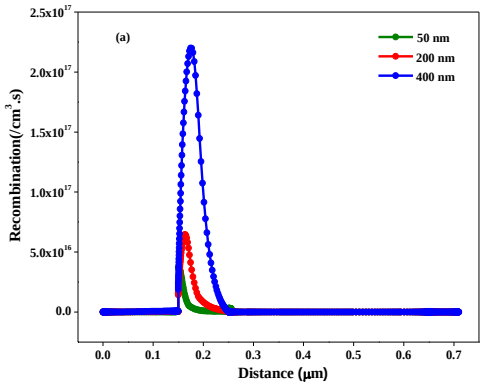


Fig. 3. Carrier recombination rate curve inside ETL at three different thickness values

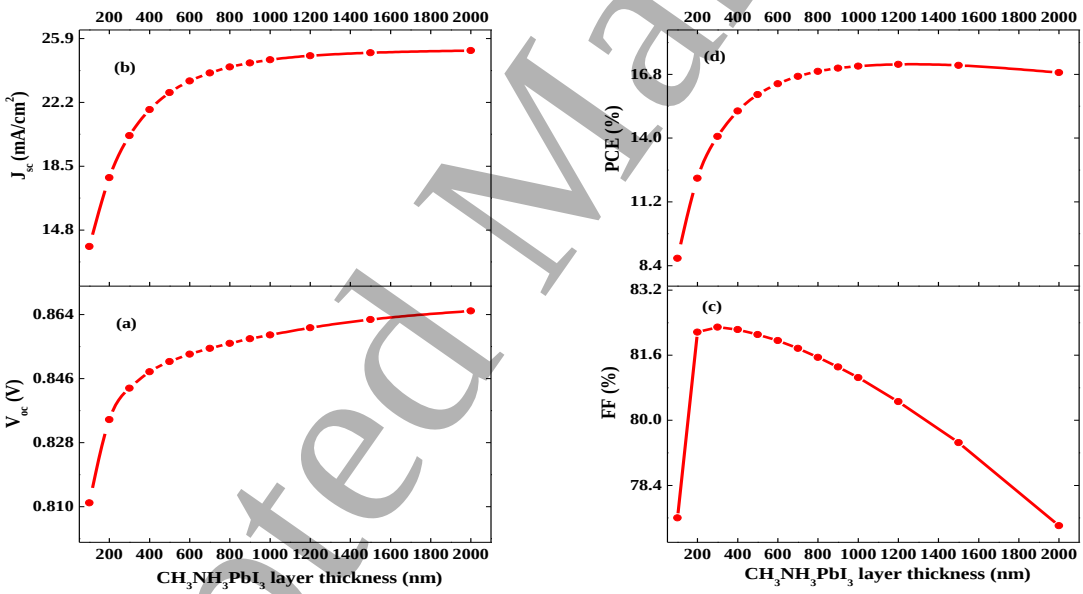


Fig.4. Change in (a) V_{oc}, (b) J_{sc}, (c) FF, and (d) PCE with the absorber layer thickness of Pb-based PSCs. Thickness of the ETL and HTL layer is 50 nm and 150 nm, respectively.

Validation of results:

To validate our theoretical results with the experimental results, we kept the ETL, perovskite, and HTL thickness at values of 30 nm, 610 nm, and 136 nm, respectively, as mentioned in the paper [29]. The obtained results have been tabulated in Table 3. Thus, the presented simulation results are in good agreement with the experimental results [29].

Accepted Manuscript

Table3: Performance parameters of both simulated and experimental MAPI-based solar cell

Structure	V_{oc} (V)	J_{sc} (mA/cm^2)	FF (%)	PCE (%)
SnO_2/$CH_3NH_3PbI_3$/CIS QD (simulation results)	0.8422	23.49	82.60	16.35
SnO_2/$CH_3NH_3PbI_3$/CIS QD (experimental results)	1.073	23.5	74.5	18.8

3.2 Performance of CIS QD based PSCs with different absorbers and ETLs

To explore the application of CIS QD HTL in PSCs, we simulated twenty PSC structures with different absorbers and ETLs. The performance of the simulated cell structures is optimized by varying the ETL and absorber thicknesses, keeping all other material parameters at fixed values. The thickness values for ETLs and absorber layers are varied from 50-400 nm and 100-1000 nm respectively. The PCE of various simulated structures is enhanced with the absorber thickness and decreases with an increase in ETL thickness. All the investigated cell structures show maximum PCE at 50 nm and 1000 nm for the ETLs and the absorbers, respectively. As seen in table 4, the cell structure with SnO_2 ETL is more efficient than PSCs structures with other ETLs. The cell parameter variations of SnO_2-based PSC with the different absorber layers thicknesses are depicted in Fig.5.

From Fig. 5, it is clear that the MASI- based cell with SnO_2 and IGZO ETLs is of comparable performance and are more efficient as compared to other PSCs structures. SnO_2 has some advantages over IGZO such as high electron mobility, antireflection property and more deeper conduction band as compared to IGZO. This may increase the transport of electrons from the perovskite absorber to the ETL and decrease the accumulation of charge carriers at the interface. Therefore, many research groups have widely used SnO_2 as an ETL [46, 47]. Furthermore, the Shockley−Quisser limit predicts that small band gap materials are best candidates to enhance the efficiency of PSCs. It is to be noted that MASI- based cell is most promising one due to its lower band gap (1.3 eV) and less toxic character as compared to the other lead-based perovskite absorbers used in the present study. Due to this, the MASI based cell has been selected with SnO_2 as ETL among all the PSCs structures. The PCE, J_{sc}, FF, and V_{oc} of the cell have values of 21.79 %, 34.86 mA/cm^2, 83.65 %, and 747.20 mV, respectively. The high value of J_{sc} is also due to the small but finite values of R_s that were taken into consideration while optimizing the device structures. Refering to Fig. 9, it can be seen that as the series resistance is increased, the value of J_{sc} decreases approaching the range which is usually obtained experimentally.

Accepted Manuscript

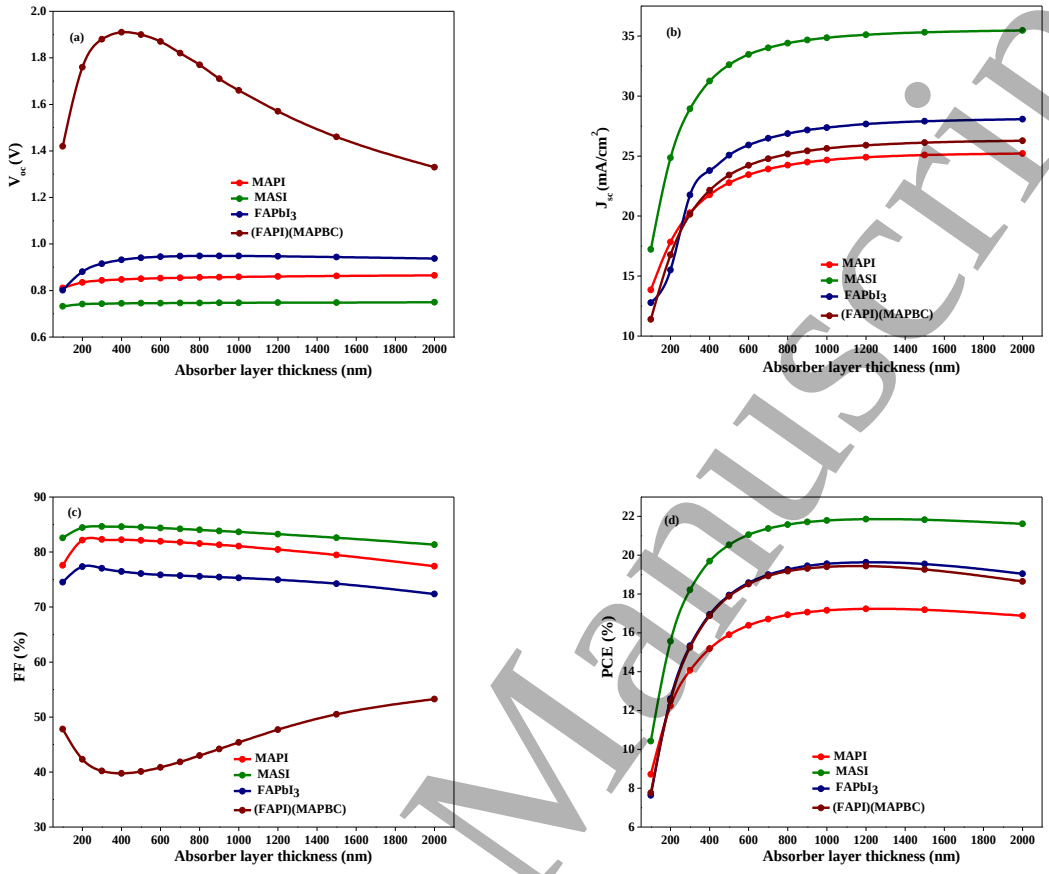


Fig.5. Performance of PSCs with SnO_2 ETL and different absorber layers.Variation in (a) V_{oc}, (b)J_{sc}, (c) FF, and (d) PCE with the absorber thickness.

Table 4. Optimized performance parameters of different absorber layer with different ETLs

Device structure	V_{oc} (V)	J_{sc} (mA/cm^2)	FF (%)	PCE (%)
$SnO_2/CH_3NH_3PbI_3$/CIS QD	0.8583	24.67	81.05	17.16
$TiO_2/CH_3NH_3PbI_3$/CIS QD	0.8584	24.67	80.86	17.13
ZnO/$CH_3NH_3PbI_3$/CIS QD	0.8589	24.67	80.30	17.02
$C_{60}/CH_3NH_3PbI_3$/CIS QD	0.8586	24.33	81.20	16.69
IGZO/$CH_3NH_3PbI_3$/CIS QD	0.8586	24.67	80.72	17.10
$SnO_2/CH_3NH_3SnI_3$/CIS QD	0.7472	34.86	83.65	21.79
$TiO_2/CH_3NH_3SnI_3$/CIS QD	0.7472	34.86	83.62	21.78
ZnO/$CH_3NH_3SnI_3$/CIS QD	0.7472	34.86	83.62	21.78
$C_{60}/CH_3NH_3SnI_3$/CIS QD	0.7471	34.57	83.68	21.61
IGZO/$CH_3NH_3SnI_3$/CIS QD	0.7472	34.86	83.64	21.79
$SnO_2/FAPbI_3$/CIS QD	0.9486	27.37	75.29	19.56
$TiO_2/FAPbI_3$/CIS QD	0.9479	27.38	75.30	19.54
ZnO/$FAPbI_3$/CIS QD	0.9457	27.38	75.34	19.51
$C_{60}/FAPbI_3$/CIS QD	0.9498	27.06	75.27	19.35

Accepted Manuscript

IGZO/$FAPbI_3$/CIS QD	0.9473	27.38	75.31	19.54
SnO_2/(FAPI)(MAPBC) /CISQD	1.66	25.64	45.38	19.40
TiO_2/(FAPI)(MAPBC)/ CIS QD	1.56	25.64	48.13	19.27
ZnO/(FAPI)(MAPBC)/ CIS QD	1.34	25.64	54.39	18.77
C_{60}/(FAPI)(MAPBC)/ CIS QD	1.80	25.28	42.29	19.28
IGZO/(FAPI)(MAPBC)/ CIS QD	1.49	25.64	49.98	19.14

Effect of defect density and the interface defects:

The defect density affects the quality of the absorber layer and the interface quality between ETL/absorber and absorber/HTL, due to which the performance of the cell gets affected. The defects enhance the charge carrier recombination and hence reduce the lifetime of the charge carrier.

From the above results, the MASI-based PSC has been found to have better cell structure; therefore, we investigated the dependence of the device parameters on the defect density (N_t) of the absorber layer. The value of N_t is varied from 1×10^{11} cm^{-3} to 1×10^{17} cm^{-3} in the MASI absorber, and the behavior of cell parameters with N_t are depicted in Fig.6. The lower values of N_t, have been found to significantly improve the cell's performance. When the N_t is kept as low as 1×10^{12} cm^{-3} in the absorber, the PCE of the cell was obtained to be 22.13 %. Further increase in the value of N_t (for example, 1×10^{15} cm^{-3}or above), leads to lower carrier diffusion length (L_p or L_n), and the cell performance shows a rapid decrement.

Accepted Manuscript

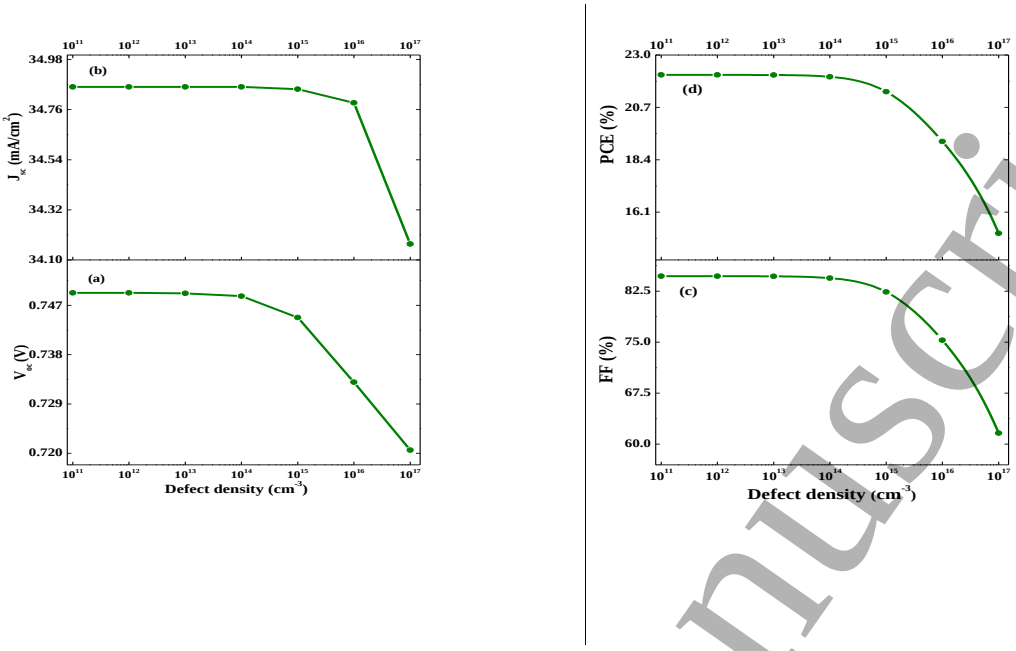


Fig.6. Effect of defect density on (a) V_{oc}, (b) J_{sc}, (c) FF, and (d) PCE of MASI based perovskite solar cell.

The device performance gets greatly affected by the defect density at the interfaces between ETL/absorber and absorber/HTL. In a real device, it is difficult to quantify the effect of recombination rate and interface defect density on cell performance. Therefore, to analyze the significance of the interface quality on overall cell performance, we varied the defect density at both the interfaces from 1×10^{9} cm^{-2} to 1×10^{15} cm^{-2}. Response of the cell parameters with interface defect density is shown in Fig.7.

It is clear that the defect density at the ETL/absorber interface has a stronger impact as compared to the absorber/HTL interface. This effect is attributed to the n-i-p structure of PSC because the light is incident from the ETL side. With the increase in defect density at the ETL/absorber interface, the generated electron-hole pairs encounter more trapping centers. Hence, charge carriers have more recombination and show non-availability for further conduction. The carrier density is more on the light irradiated side than on the backside. Thus, the significant carrier loss due to recombination is observed more at the ETL/absorber interface than the absorber/HTL interface. This may be due to the collection probability of photogenerated charge carries. If we reverse the cell structure i.e. p-i-n structure, the defect density is influenced more at the HTL/absorber interface as compared to the absorber/ETL interface. This is again due to an increase in trapping centre at the HTL/absorber interface with the increase in defect density as compared to the back side.

Hence, in order to increase the performance of the PSC, improved interface quality at the front side plays an important role rather than at the back side. A similar effect has also been observed

Accepted Manuscript

by another researcher [48]. For this reason, a high-quality ETL/absorber interface is important for highly efficient n-i-p PSC.

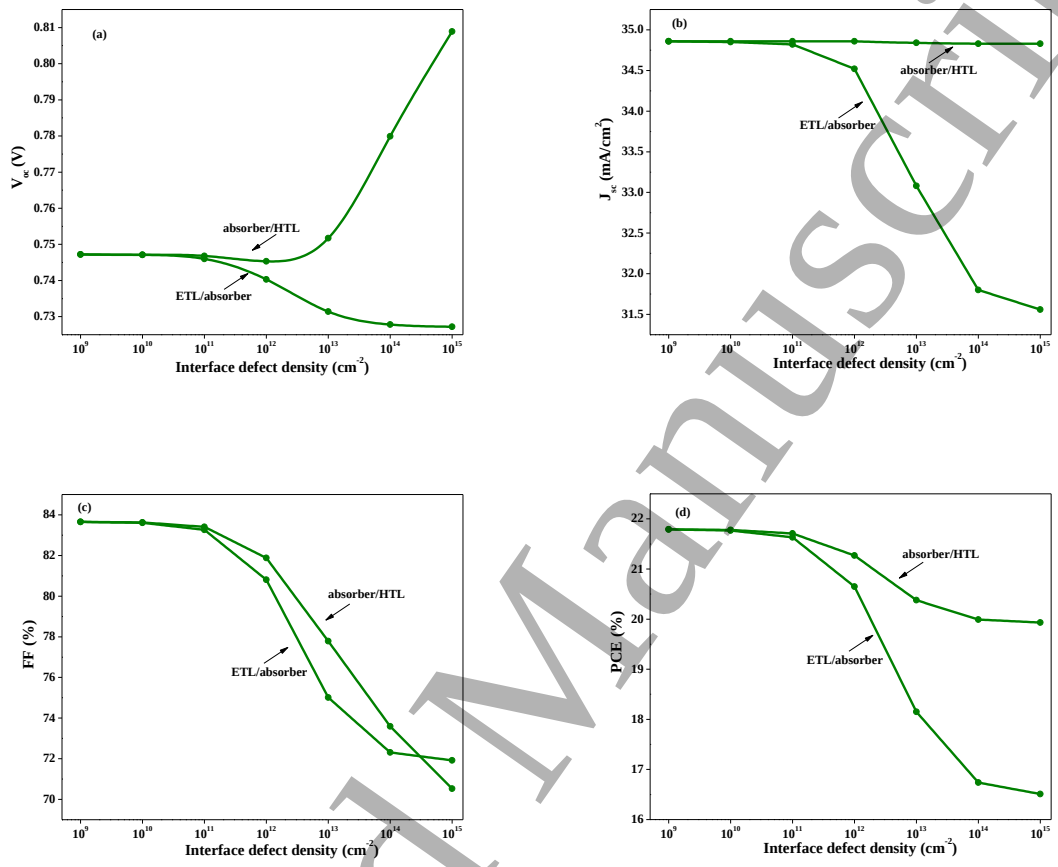


Fig.7. Effect of defect densities of ETL/absorber and absorber/HTL interfaces on (a) V_{OC}, (b) J_{SC}, (c) FF, and (d) PCE of perovskite solar cell.

Effect of metal work function: The work function of the metal electrode connected as the back contact in PSC plays a vital role in charge carrier extraction. To observe the hole extraction rate of various metals such as Cr (4.5 eV), Graphene (4.6 eV), Cu (4.65 eV), Ag (4.7 eV), C (5.0 eV), Au (5.1eV), and Ni (5.15 eV) in PSC, we have used each metal one by one in $MASnI_3$ based PSC as back contact. The cell results with the above-mentioned metals are shown in Fig.8. Higher metal work function results in higher charge extraction which, in turn, improves the cell performance. Therefore, Au and Ni yield encouraging results as compared to other metals due to their high work function. The high work function of Au and Ni also creates a greater barrier for electron transfer from HTM to metal [49].

Accepted Manuscript

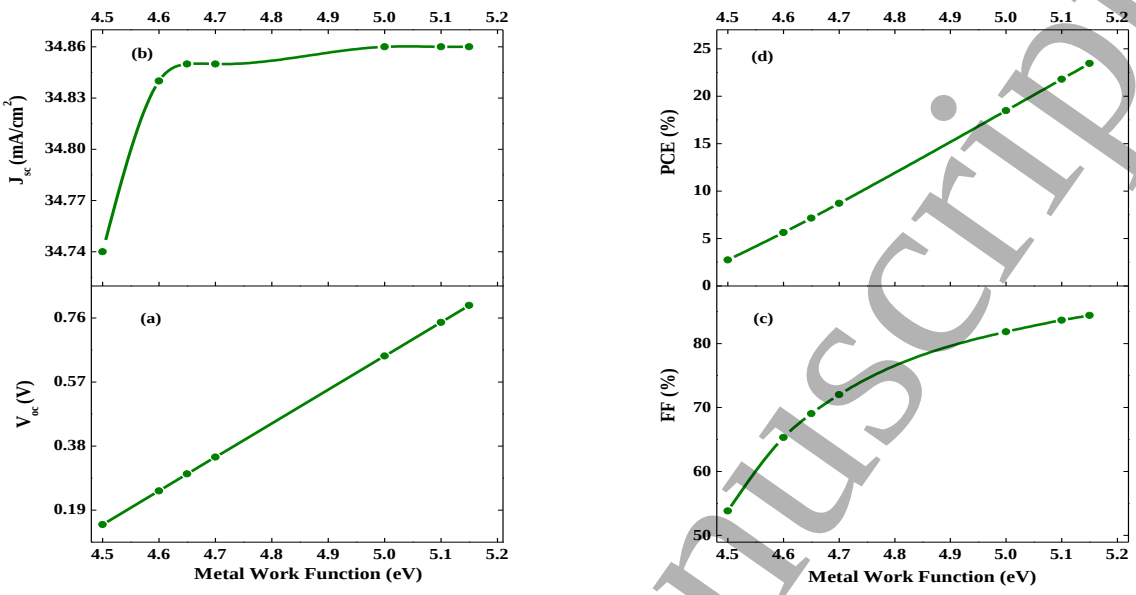


Fig. 8. Change in (a) V_{oc}, (b) J_{sc}, (c) FF, and (d) PCE with the metal work function

Effect of R_s and R_p:

The series (R_s) and parallel resistance (R_p) critically impact the performance of the cell. Series resistance is based on electrical resistance associated with the front and back contacts (ITO and Au) and electrical dissipation in the ETL, absorber, and HTL layers. Parallel resistance gets mainly affected by the design of the device, which indicates various alternative recombination paths. As a result, the low value of parallel resistance affects the collected photocurrent and causes a photovoltage loss. Meanwhile, series resistance mainly affects the J_{sc}, FF, and PCE values.

The effects of both the resistances have been observed and analyzed on cell performance. The R_s and R_p resistance values are varied in the range of 5-100 ohm-cm^2 and 500-5000 ohm-cm^2, respectively. The effect of series and parallel resistance on the performance parameters of the cell is shown in Fig. 9 and Fig. 10, respectively. In Fig. 9, it is observed that the increase in R_s strongly reduces the cell parameters. The J_{sc} decreases from 34.86 mA/cm^2 to 7.39 mA/cm^2, FF decreases from 62.70 % to 18.82 %, and PCE decreases from 16.35 % to 1.04 %, but V_{oc} behavior is almost unaffected.

Fig. 10 depicts that with the rise in R_p, the cell parameters increase in their values except for J_{sc}. Therefore, a better cell structure has a low value of R_s and a high value of R_p. Our simulation results are in good agreement with the results reported in the literature [6].

Accepted Manuscript

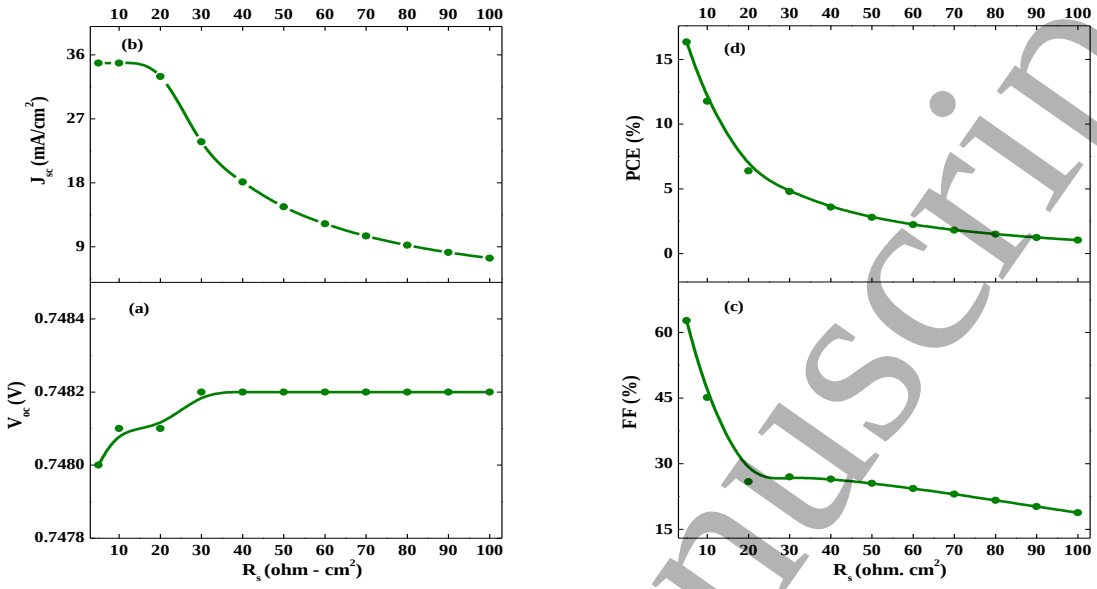


Fig. 9. Change in (a) V_{oc}, (b) J_{sc}, (c) FF, and (d) PCE with R_s

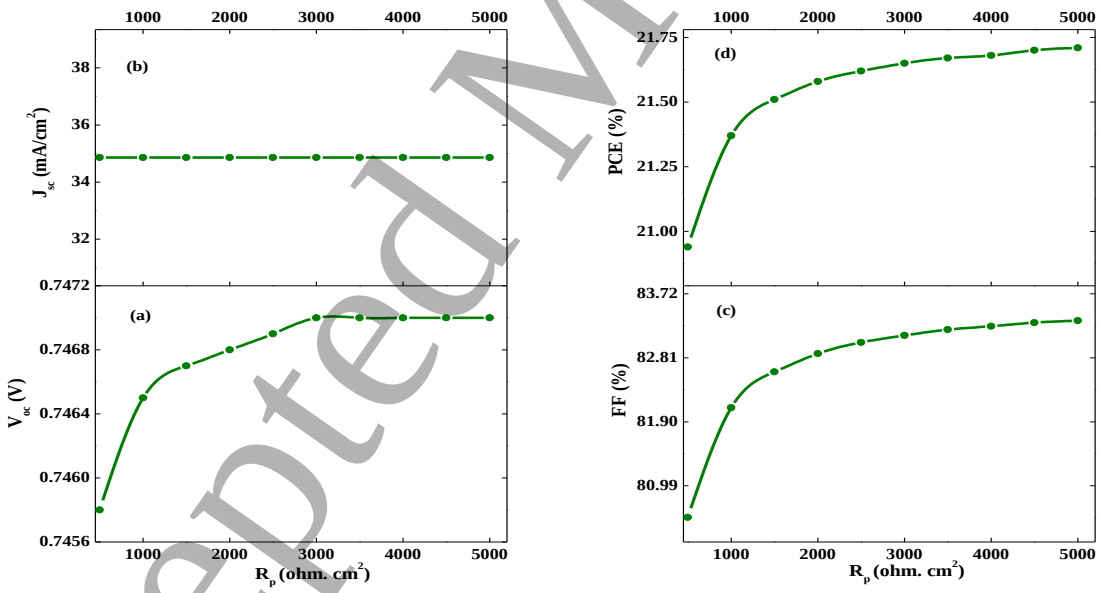


Fig. 10. Change in (a) V_{oc}, (b) J_{sc}, (c) FF, and (d) PCE with R_p

Effect of Cell Temperature: The operating temperature is varied from 300-500 K to observe its effect on the solar cell performance. The cell parameter variation with temperature is shown in Fig. 11. The increase in cell temperature reduces the diffusion length and increases the series resistance, resulting in reduced cell performance. The cell performance is found to be optimum at 300 K, where J_{sc}= 34.86 mA/cm^2, V_{oc} = 747.20 mV, FF = 83.86 % and PCE = 21.79 %. Hence,

Accepted Manuscript

we carried out the analysis of the MASI based PSC at a 300 K temperature to realize an efficient performance.

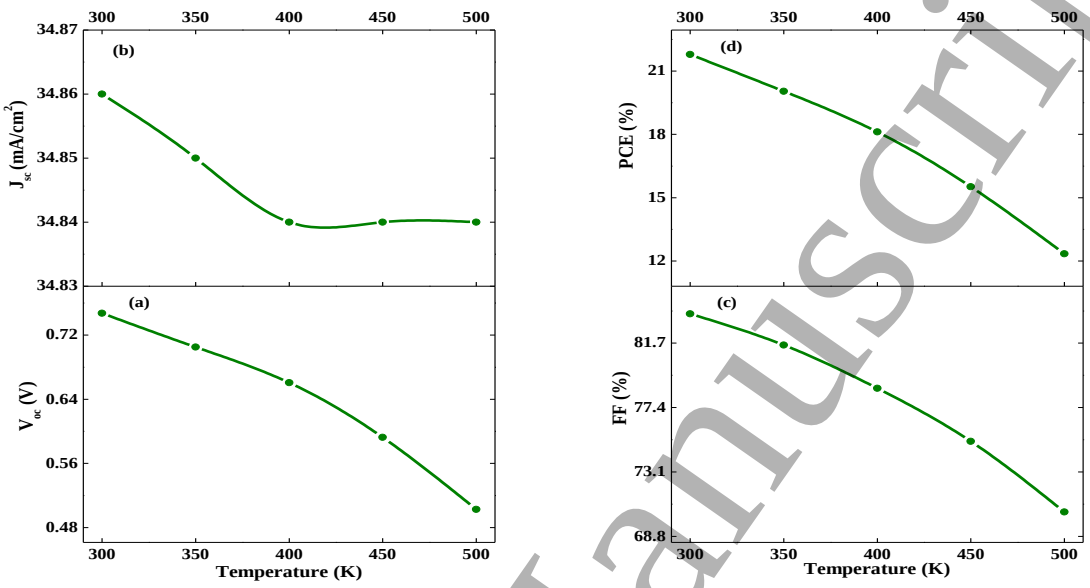


Fig. 11. Change in (a) V_{oc}, (b) J_{sc}, (c) FF, and (d) PCE with temperature

Conclusion: Simulation studies for the experimentally analyzed perovskite solar cell, i.e., ITO/SnO_2/$CH_3NH_3PbI_3$/$CuInS_2$/Au, have been carried out using SCAPS-1D software. The results of our study are in good agreement with the experimental results. The use of inorganic $CuInS_2$ quantum dots in various perovskite solar cell structures has been explored. The highest performance (PCE = 21.79 %) has been displayed by the $MASnI_3$ based PSC structure as compared to other mentioned cells, i.e., ITO/SnO_2/$CH_3NH_3SnI_3$/$CuInS_2$/Au. The increase in defect density in the absorber layer reduced the overall performance of the cell. The interfacial defect density has been found to have more impact on the ETL/absorber interface as compared to the absorber/HTL interface. The metals Au and Ni give better performance as compared to other used metals as back contact. The effect of series and parallel resistances has a good impact on cell performance. The rise in cell operating temperature has also been observed to produce a detrimental effect on the cell parameters. The present work is useful in framing an efficient MASI based perovskite solar cell with CIS QD.

Acknowledgment

All the authors gratefully acknowledge the Prof Marc Burgelman, Department of Electronics and Information System at the University of Gent Belgium for providing the access of SCAPS software.

Accepted Manuscript

References:

[1]. Fakharuddin, A., De Rossi, F., Watson, T. M., Mende, L. S., and Jose, R., 2016. Research Update: Behind the high efficiency of hybrid perovskite solar cells. APL MATERIALS. 4, 091505.

[2]. Sahu, A., and Dixit, A., 2018. Inverted structure perovskite solar cells: A theoretical study. Current Applied Physics. 18, 1583-1591.

[3] Kojima, A., Teshima, K., Shirai, Y., Miyasaka, T., 2009. Organometal Halide Perovskites as Visible-Light Sensitizers for Photovoltaic Cells. J. Am. Chem. Soc. 131, 6050.

[4] Duan, Q., Ji, J., Hong, X., Fu, Y., Wang, C., Zhou, K., Liu, X., Yang, H., Wang Z.-Y., 2020. Design of hole transport-material free CH3NH3PbI3/CsSnI3 all-perovskite heterojunction efficient solar cells by device simulation. Solar Energy. 201, 555–560.

[5]. Jiang, X., et al., 2021. Tin Halide Perovskite Solar Cells: An Emerging Thin-Film Photovoltaic Technology. Acc. Mater. Res.

[6]. Karthicka, S., Velumania, S., and Bouclé, J., 2020. Experimental and SCAPS simulated formamidinium perovskite solar cells: A comparison of device performance. Solar Energy. 205, 349–357.

[7]. Eperon, G.E., Stranks, S.D., Menelaou, C., Johnston, M.B., Herz, L.M., Snaith, H.J., 2014. Formamidinium lead trihalide: A broadly tunable perovskite for efficient planar heterojunction solar cells. Energy Environ. Sci. https://doi.org/10.1039/ c3ee43822h.

[8]. Chen,Y., Hu,Y., Meng, Q., Yan, H., Shuai, W., and Zhang, Z., 2019. Natively textured surface of Ga-dopedZnO films electron transporting layer for perovskite solar cells: further performance analysis from device simulation, Journal of Materials Science: Materials in Electronics. 30, 4726–4736.

[9]. Wang, P., et al. 2020. Ion Exchange/Insertion Reactions for Fabrication of Efficient Methylammonium Tin Iodide Perovskite Solar Cells. Adv. Sci.7, 1903047.

[10]. Gagandeep, Singh, M., Kumar, R., Singh, V., 2019. Graphene as charge transport layers in lead free perovskite solar cell. Mater. Res. Express. 6, 115611.

[11]. Jiang, X., Wang, F., Wei, Q., Li, H., Shang, Y., Zhou, W., Wang, C., Cheng, P., Chen, Q., Chen, L., & Ning, Z., 2020. Ultra-high open-circuit voltage of tin perovskite solar cells via an electron transporting layer design. nature communications. 11, 1245.

[12]. Noel, N. K., et al 2014. Lead-free organic–inorganic tin halide perovskites for photovoltaic applications. Energy Environ. Sci. 9, 3061.

[13]. Baig, F., Khattak, Y. H., Mari, B., Beg, S., Ahmed, A., and Khan, K., (2018), Efficiency Enhancement of $CH_3NH_3SnI_3$ Solar Cells by DeviceModeling. Journal of Electronic Materials 47, 9.

[14]. Ke,W., Fang, G., Liu, Q., et al., 2015. Low-Temperature Solution-Processed Tin Oxide as an Alternative Electron Transporting Layer for Efficient Perovskite Solar Cells. J. Am. Chem. Soc. 137, 6730–6733.

[15]. Liu, D., and . Kelly, T. L., 2013. Perovskite solar cells with a planar heterojunction structure prepared using room-temperature solution processing techniques, Nat. Photonics. 8, 133–138.

[16]. Azri, F., Labed, M., Meftah, A.F., Sengouga, N., Meftah, A.M., 2016. Optical characterization of aIGZO thin film for simulation of a-IGZO(n)/µ-Si(p) heterojunction solar cell. Opt. Quant. Electron. 48, 391.

[17]. Leeb , K.-M., Chena ,C.-C., et al., 2017. Thickness effects of thermally evaporated C60 thin films

Accepted Manuscript

on regular-type CH3NH3PbI3 based solar cells. Solar Energy Materials & Solar Cells. 164, 13–18.
[18]. Iftiquar, S.M., Kim,J.S., Yi, J., 2017. Investigation of highly efficient methyl ammonium lead halide perovskite solar cell with non-textured front surface. Optik 148, 54–62.
[19].Zhou, H., et al., 2014. Photovoltaics interface engineering of highly efficient perovskite solar cells, Science. 345, 542–546.
[20] Xiao,Y., et al., 2014. An all-solid-state perovskite-sensitized solar cell based on the dual function polyaniline as the sensitizer and p-type hole-transporting material, J. Power Sources. 267, 1–8.
[21] Endres J.,Kulbak M., Zhao L., Rand B.P., CahenD., HodesG., KahnA., Electronic structure of the CsPbBr3/polytriarylamine (PTAA) system, (2017), J. Appl. Phys. 121 035304.
[22] QinP., TanakaS., Ito S., TetreaultN., ManabeK., Nishino H.,NazeeruddinM.K., GrätzelM., (2014), Inorganic hole conductor-based lead halide perovskite solar cells with 12.4% conversion efficiency, Nat. Commun. 5, 3834–3839.
[23] YangY., Chen H., ZhengX., MengX., ZhangT., HuC., BaiY., XiaoS., YangS., (2017) Ultrasound-spraydeposition of multi-walled carbon nanotubes on NiO nanoparticles-embedded perovskite layers for high-performance carbon-based perovskite solar cells, Nano Energy ,42, 322–333.
[24] ZhangY. et. al., (2019) An inorganic holetransport materialof CuInSe2 for stable and efficient perovskite solar cells, Org. Electron. 67 168–174.
[25] Lv, M. et. al, (2015), Colloidal CuInS2 quantum dots as inorganic hole-transporting material in perovskite solar cells, ACS Appl. Mater. Interfaces, 7, 17482–17488.
[26] Wu, Q., et. al., (2015), Kesterite Cu_2ZnSnS_4 as a low-cost inorganic hole-transporting material for high-efficiency perovskite solar cells, ACS Appl. Mater. Interfaces, 7, 28466–28473.
[27] Xu, L., et. al., (2017), Solution-processed Cu(In, Ga)(S, Se)$_2$nanocrystal as inorganic hole-transporting material for efficient and stable perovskite solar cells, Nanoscale Res. Lett., 12, 159.
[28] Kolny, O. J., Weller, H. (2013), Synthesis and application of colloidal $CuInS_2$ semiconductor Nanocrystals ACS Appl. Mater. Interfaces 5, 12221–12237.
[29]. Liua Y. et. al., (2019). A novel inorganic hole-transportingmaterial of $CuInS_2$ for perovskite solar cells with high efficiency and improved stability, Organic Electronics 75, 105430.
[30] Burgelman, M., Nollet, P., Degrave, S., (2000), Modelling polycrystalline semiconductor solar cells. Thin Solid Films, 361-362, 527- 532.
[31]. Minemotoa, T., and Murata, M., (2014), Impact of Work Function of Back Contact of Perovskite Solar Cells without Hole Transport Material Analyzed by Device Simulation, Current Applied Physics, 181, 372-378.
[32].Ma, F., Li, J., Li, W., Lin, N., Wang, L., Qiao, J., (2017) Stable α/δ phase junction of formamidiniumlead iodide perovskites for enhanced near-infrared emission. Chem. Sci., **8**, 800-805.
[33]. Chen, L.C., Tseng, Z.L., Huang, J.K., (2016), A study of inverted-type perovskite solar cells with various composition ratios of (FAPBI3)1–x(MAPbBr 3)x. Nanomaterials 6, 183
[34]. Gélvez-Rueda, M.C., Renaud, N., Grozema, F.C., (2017) Temperature dependent charge carrier dynamics in formamidinium lead iodide perovskite. J. Phys. Chem. C, 12142, 23392–23397.
[35]. Kaifi, M. and GuptaS. K., (2019),Simulation of Perovskite based Solar Cell and Photodetector using SCAPS Software, International Journal of Engineering Research and Technology. ISSN 0974-3154,12, 1778-1786.
-[36]. Mandadapu, Usha, VedanayakamVictor S., ThyagarajanK.and BabuB.J., (2018). Optimization of high efficiency tin halide perovskite solar cells using SCAPS-1D , Int. J. Simulation and Process Modelling, 13,3.

Accepted Manuscript

[37]. Hirasawa, M., Ishihara, T. et. al.,(1994). Magnetoabsorption of the lowest exciton in perovskite-type compound CH3NH3PbI3. Physica B, 201, 427..

[38]. Liu, M., Johnston, M. B., Snaith, H. J., (2013), Efficient planar heterojunction perovskite solar cells by vapor deposition. Nature 501, 395.

[39]. Lee, M. M. et. al., 2012. Efficient hybrid solar cells based on meso-super structure organometal halide perovskites. Science, 338, 643.

[40]. Yue, W. et. al., 2010.$CuInS_2$ quantum dots synthesized by a solvothermal route and their application as effective electron acceptors for hybridsolar cells, J. Mater. Chem. 20, 7570–7578.

[41]. Karimi, E., and Ghorashi, S. M. B., (2019), The Effect of SnO_2 and ZnO on the Performance of Perovskite Solar Cells. Journal of Electronic materials 49, 364-376.

[42] Azri F., Meftah, A., Sengouga, N., Meftah, A., 2019. Electron and hole transport layers optimization by numerical simulation of a perovskite solar cell, Solar Energy, 181, 372–378.

[43].Minemoto, T., and Murata, M., 2014. Device modeling of perovskite solar cells based on structural similarity with thin film inorganic semiconductor solar cells. Journal of Applied Physics 116, 054505.

[44] Raoui, Y., Zahraouy, H. E., Kazim, S., Ahmad, S., (2020), Energy Level Engineering of Charge Selective Contact and Halide Perovskite by Modulating Band Offset: Mechanistic Insights, ChemRxiv.

[45] Bag, A., Radhakrishnan, R., Nekovei, R., and Jevakumar, R., 2020, Effect of absorber layer, hole transport layer thicknesses, and its doping density on the performance of perovskite solar cells by device simulation. Solar Energy 196, 177–182.

[46]. Zhao, P. et. al. ,2019. Numerical Simulation of Planar Heterojunction Perovskite Solar Cells Based on SnO_2 Electron Transport Layer, ACS Appl. Energy Mater., 2, 4504−4512.

[47]. Jiang, Q., Zhang, X., and You, J., 2018. SnO_2: A Wonderful Electron Transport Layer for PerovskiteSolar Cells, Small, 1801154.

[48]. Huang, S., Rui, Z., Chi, D., and Bao, D. 2019. Influence of defect states on the performances of planar tinhalide perovskite solar cells, J. Semicond. 40 (3), 032201.

[49]. Behrouznejad et. al., 2016. A study on utilizing different metals as the back contact of $CH_3NH_3PbI_s$ perovskite solar cells, J. Mater. Chem. A, 4, 13488.

Accepted Manuscript