MEDICAL IMAGE ANALYSIS FOR DIAGNOSING AND CLASSIFYING BREAST CANCER USING DEEP LEARNING

Thesis Submitted
in Partial Fulfillment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

by

GUNJAN CHUGH
(2k18/PHD/CO/22)

Under the Joint Supervision of

Prof. Shailender Kumar (Supervisor)
Professor, Department of CSE
DTU, Delhi

Prof. Nanhay Singh (Co-Supervisor)
Professor, Department of CSE
NSUT (East Campus), Delhi

Department of Computer Science and Engineering

DELHI TECHNOLOGICAL UNIVERSITY, DELHI

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-110042, India

July, 2024

ACKNOWLEDGEMENT

I express my heartfelt gratitude to Almighty God for providing me with the strength to complete my dissertation and guiding me throughout my life.

It gives me great pleasure to express my sincere gratitude to everyone who has assisted me with this research. It gives me immense pleasure to convey my sincere appreciation and obligation of thanks to my knowledgeable superiors **Prof. Shailender Kumar**, Dept. of Computer Science & Engineering, Delhi Technological University, New Delhi, and **Prof. Nanhay Singh**, Dept. of Computer Science & Engineering, Netaji Subhas University of Technology, New Delhi for their invaluable guidance, encouragement, and patient reviews. I am incredibly grateful to both of my supervisors for helping me with the research work and providing me with moral support throughout the course.

I am extremely thankful to **Prof. Vinod Kumar**, Head of the Department, Department of Computer Science and Engineering, and **Prof. Rahul Katarya,** Chairperson DRC, Dept. of Computer Science & Engineering for their insightful comments and valuable suggestions. I am also thankful to all teaching and non-teaching staff at DTU, who have helped me directly or indirectly complete this research plan. Also, I sincerely thank Delhi Technological University for considering my candidature for this course.

Most prominently, I am deeply grateful to my parents **Dr. S. P. Chugh** and **Prof. Promila Chugh,** and in-laws, **Mr. S.K. Pruthi** and **Mrs. Sushma Pruthi** who have encouraged and motivated me, and without their support, I could not have completed this work. I am extremely thankful to my husband **Mr. Ankur Pruthi** and my daughter **Ms. Namya Pruthi** for their moral and emotional support in my life. Without their tremendous support, understanding and encouragement over the past few years, it would have been impossible for me to complete my study.

Gunjan Chugh

Roll no: 2K18/PhD/CO/22

Department of Computer Science

Delhi Technological University, Delhi, India-110042

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahabad Daulatpur, Main Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, Gunjan Chugh(Roll no: 2K18/PhD/CO/22) hereby certify that the work which is being presented in the thesis entitled **"Medical Image Analysis for Diagnosing & Classifying Breast Cancer using Deep Learning"** in partial fulfillment of the requirements for the award of the Degree of Doctor of Philosophy, submitted in the **Department of Computer Science & Engineering**, Delhi Technological University is an authentic record of my own work carried out during the period from July 2018 to July 2024 under the supervision of **Prof. Shailender Kumar (Supervisor)**, Department of Computer Science & Engineering, DTU, Delhi and **Prof. Nanhay Singh (Co-Supervisor)**, Department of Computer Science & Engineering, NSUT (East Campus), Delhi.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other Institute.

Candidate's Signature

This is to certify that the student has incorporated all the corrections suggested by the examiners in the thesis and the statement made by the candidate is correct to the best of our knowledge.

Signature of Supervisor(s) Signature of External Examiner

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahabad Daulatpur, Bawana Road, Delhi-110042

CERTIFICATE BY THE SUPERVISOR(s)

Certified that **Gunjan Chugh (Roll no: 2K18/PhD/CO/22)** has carried out her research work presented in this thesis entitled "**Medical Image Analysis for Diagnosing & Classifying Breast Cancer using Deep Learning**" for the award of **Doctor of Philosophy** from the Department of Computer Science and Engineering, Delhi Technological University, Delhi under our supervision. The thesis embodies the results of original work carried out by the student herself and the contents of the thesis do not form the basis for the award of any other degree to the candidate or anybody else from this or any other University/Institution.

Prof. Shailender Kumar (Supervisor)
Professor, Department of CSE
DTU, Delhi-42

Prof. Nanhay Singh (Co-Supervisor)
Professor, Department of CSE
NSUT (East Campus), Delhi-78

ABSTRACT

Breast carcinoma is the premier category of deadliest cancer reported in females. Breast cancer cases are rising dramatically both in India and around the world, especially in individuals between the ages of 30 and 40. Automatic diagnosis of breast cancer is necessary because manual diagnosis is laborious and time-consuming. Our current healthcare systems are prone to failure. Late detection is the main cause of low breast carcinoma survival rates in the country. Consequently, computer-aided diagnosis (CAD) for medical imaging has become a useful gadget for physicians to categorize clinical images into several groups, facilitating early diagnosis and treatment. Machine Learning (ML) and Deep Learning (DL) have developed various techniques/algorithms for diagnosing and classifying breast cancer early. Multiple strategies have been employed by experts to anticipate health issues before they manifest symptoms. Consequently, in the medical and healthcare communities, getting a precise diagnosis and prognosis of tumors is considered a difficult endeavour for doctors.

This research thus addresses the need for medical image analysis using CAD for early diagnosis and prognosis in the healthcare domain. The literature survey communicated cutting-edge research disseminated in breast malignancy using ML and DL approaches. Although malignancy can't be proven without biopsy, early carcinoma detection using imaging modalities is an hour of need. Mammography is employed as the "benchmark" for breast carcinoma examination, owing to its widespread availability and cost-effectiveness compared to others. Current research limitations suggest that technical and practical investigation is desperately needed to boost healthcare over the long term.

'Transnet' is the first CAD model proposed in this research to diagnose and classify breast carcinoma with enhanced performance. Two experiments were performed on the Curated Breast Imaging Subset-Digital Database of Screening Mammography (CBIS-DDSM) dataset. The following Deep Neural Networks(DNN) were utilized- VGG-16, VGG-19, Mobile Net, ResNet-50, ResNet-152, and DenseNet-169. In the first experiment, namely Deep feature fusion with ML Classifier, pre-trained networks

were deployed as feature extractors, and afterward, the acquired attributes were provided to machine learning classifiers for classification. The second experiment, called Deep feature fusion with Neural Net classifiers, fine-tuned these networks for feature extraction and categorization. The findings revealed that the proposed approach performed remarkably well than the other cutting-edge methodologies. The second approach performed better than the first, thus, improving all the evaluation metrics.

Another CAD framework proposed to enhance performance through smaller datasets is the Multi Stage Transfer Learning Approach (MSTLA). Three mammography datasets were utilized: Mammography Image Analysis Society (MIAS), In-Breast, and CBIS-DDSM. The model was fine-tuned in three stages on separate datasets, and the optimized DCNN was carried forward at the next stage. Two DNNs were deployed for training the model – DenseNet-169 and ResNet-152. The results have shown that Stage 3 performs best compared to the other two stages, with DenseNet-169 having accuracy and AUC values of 100 and 1.0. Thus, the proposed approach could be employed for early-stage breast carcinoma diagnosis.

DNNs can memorize the training information owing to their huge learning capacity. In the medical domain, there is an urgent need to assess the generalizability of deep neural networks. Generalization is an approach to analyze how the model behaves on unseen data. Generalization Error(GE) measures the difference between training and testing errors. Thus to address this gap, we have proposed another framework for evaluating the generalization error in DNN. Gaussian, Salt and pepper, and Speckle noise were added to the CBIS-DDSM dataset. Generalizability was evaluated for three DCNNs - InceptionNet v3, DenseNet-201, and EfficientNet-B4. Results have shown that the proposed framework with DenseNet-201 has minimum generalization error and thus exhibits high generalizability on the unseen i.e. noisy data.

This research work successfully provides a more reliable, efficient, and optimal approach for early-stage breast cancer diagnosis and thus could be deployed in laboratories. Future perspectives of the proposed methodology include its implementation on various imaging modes such as Ultrasound, MRI, CT, etc.

LIST OF PUBLICATIONS

PUBLICATIONS IN SCIE JOURNALS

Published

- G. Chugh, S. Kumar, and N. Singh, “Survey on Machine Learning and Deep Learning Applications in Breast Cancer Diagnosis,” Cognitive Computation. 2021, doi: 10.1007/s12559-020-09813-6. **(SCIE Indexed, IF=4.3)**
- G. Chugh, S. Kumar, and N. Singh, “TransNet: a comparative study on breast carcinoma diagnosis with classical machine learning and transfer learning paradigm,” *Multimed. Tools Appl.*, no. 0123456789, 2023, doi: 10.1007/s11042-023-16938-x. **(SCIE Indexed, IF=3.0)**

Under Review

- G. Chugh, S. Kumar, and N. Singh, “Enhancing Generalizability and Robustness of Deep Neural Networks Trained on Mammograms through Adversarial Training.” 2024 *Biomedical Signal Processing and Control,* Elsevier **(SCIE Indexed, IF=4.9)**
- G. Chugh, S. Kumar, and N. Singh, “Ultra-sono: An optimized CAD framework for breast lesion detection & classification in ultrasound images”, 2024 *Health Information Science and Systems*, Springer **(SCIE Indexed, IF=4.7)**

PUBLICATIONS IN IEEE CONFERENCES

- G. Chugh, S. Kumar, and N. Singh, “MSTLA: Multi-Stage Transfer Learning Approach for Breast Carcinoma Diagnosis,” *2023 Int. Conf. Adv. Comput. Comput. Technol. InCACCT 2023*, pp. 509–514, 2023, doi: 10.1109/InCACCT57535.2023.10141697.
- G. Chugh, S. Kumar and N. Singh, "A Framework for Generalization Error Evaluation in Deep Convolutional Neural Networks," *2023 IEEE Engineering Informatics*, Melbourne, Australia, 2023, pp. 1-7, doi: 10.1109/IEEECONF58110.2023.10520375.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Ada Boost	Adaptive Boosting
AI	Artificial Intelligence
ANN	Artificial Neural Network
BCE	Binary Cross Entropy
CAD	Computer-Aided Diagnosis
CBIS-DDSM	Curated Breast Imaging Subset- Digital Database of Screening Mammography
CC	Craniocaudal
CCE	Categorical Cross Entropy
CNN	Convolutional Neural Network
DBT	Digital Breast Tomosynthesis
DCNN	Deep Convolutional Neural Network
DL	Deep Learning
DM	Digital Mammography
DNN	Deep Neural Networks
DT	Decision Tree
FFDM	Full-field Digital Mammography
GE	Generalization Error
GPU	Graphical Processing Unit
HP	Histopathology
KNN	K-Nearest Neighbour
LASSO	Least absolute shrinkage and a selection operator

LR	Logistic Regression
MIAS	Mammography Image Analysis Society
ML	Machine Learning
MLO	Medio lateral Oblique
MRI	Magnetic Resonance Imaging
MSTLA	Multi-Stage Transfer Learning Approach
NB	Naive Bayes
PCA	Principal Component Analysis
RBF	Radial Basis Function
RF	Random Forest
ROI	Region of Interest
SEER	Surveillance Epidemiology and End Result
SFM	Screen Film Mammography
SVM	Support Vector Machine
UCI	University of California Irvine
US	Ultrasound
WBCD	Wisconsin Breast Cancer Dataset
XGB	Extreme Gradient Boosting
YOLO	You Only Look Once

CHAPTER 1

INTRODUCTION

1.1 Breast Cancer

The body of an individual comprises billions of cells. These cells keep on growing and multiplying during their lifetime. The irregular growth of these cells forms a lump. This lump appears in the form of a tumor[1]. The tumor can be classified as non-cancerous called Benign tumors or cancerous called Malignant tumors. Benign tumors do not metastasize to other bodily regions and are confined to the organ in which they originated. Conversely, malignant tumors spread to further bodily organs and affect those organs as well[2]. Tumors can become deadly if they proliferate across the body and remain untreated.

The carcinoma that initiates in the tissues of the breast and spreads impulsively is referred to as breast carcinoma. The strongest threat for breast malignancy is in females. Approximately 99% of occurrences of breast malignancy originate in females, and 0.5–1% of cases occur in men [3]. Figure 1.1 shows the structure of a woman's breast. Breast carcinoma can develop in multiple locations on the breast. The three core elements of a breast are lobules, ducts, and connective tissue. Lobules are responsible for generating milk. The streams that transport milk to the nipple are referred to as ducts. Connective tissue encompasses and contains everything together. Majority of the breast malignancies initiates in the ducts or lobules[4]. The unbounded evolution of tumor cells can disperse to other robust breast tissue and transit to the arms' lymph glands. The lymphatic nodes move the tumor cells to further body parts.

1.1.1 Symptoms

Physical evidence of breast malignancy includes breast tenderness, lump in the breast area, irritation, nipple discharge besides breast milk, inverted nipple, modifications in breast size, inflammation, or blister under your arm. Warning signs associated with breast malignancy include overweight, low physical activity, excessive alcohol consumption, dense breasts, genes, early menstruation, and advanced maternal age[5]. Researchers concluded that the factors that are considered relevant for cancer at an initial stage might not contribute equally to predicting patients' survivability over the long term. For example, age is regarded as an essential factor for forecasting survivability, but its potentiality drops with time, maybe due to the empowerment of other factors such as grade, lymph nodes positive, etc. Therefore, the importance of all such elements must be considered when developing an architecture for predicting breast carcinoma survivability in the future.

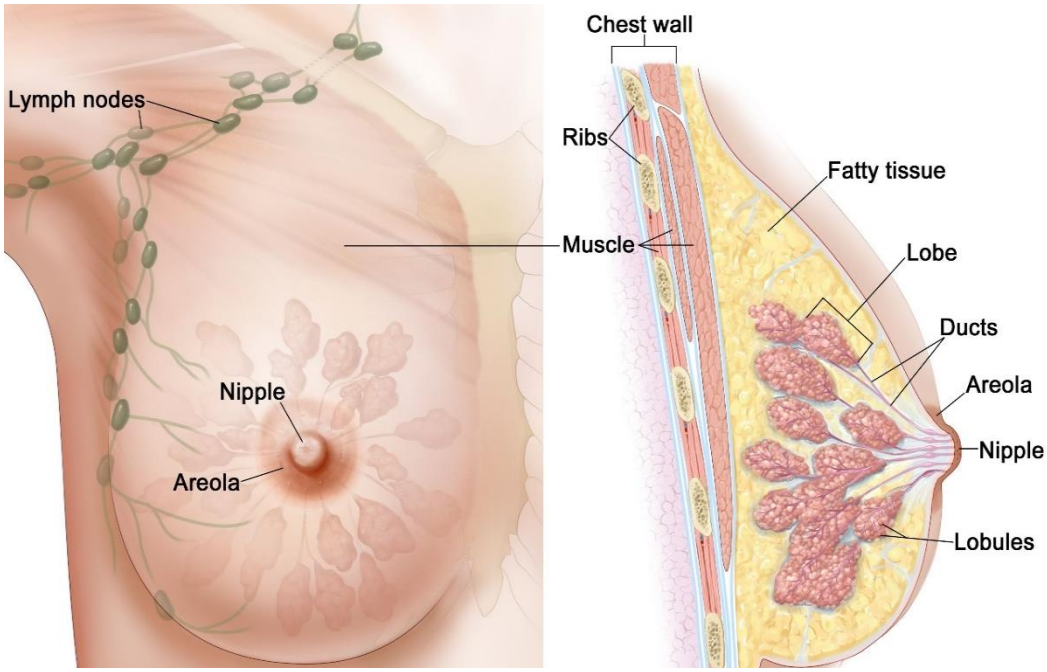

Fig.1.1: Anatomy of a Women's Breast[6]

(Source: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/breast)

1.1.2 Categorization

Breast carcinoma is investigated by distinguishing tumors as malignant and benign. Benign tumors do not metastasize to additional bodily regions and are confined to the organ in which they originated. On the contrary, malignant tumors disperse to different bodily regions and affect those organs as well. Figure 1.2 shows the two categories of tumors. With the naked eye both the mammogram images look identical. But the picture on the left-hand portion is of a benign tumor where whereas one on the right-hand side is of a malignant tumor. Thus, medical practitioners and pathologists need a reliable diagnostic process to differentiate these two classes of tumors[7].

The tumor can be further classified into two categories: Invasive Carcinoma and Non-Invasive/ In- Situ Carcinoma[7]:

(i) **Invasive cancer:** This carcinoma starts in the milk tubes(canals) and can disperse to the adjacent cells and different body parts as well.

(ii) **Non-Invasive/In-situ cancer:** This category of cancer cannot escalate within or to other body parts and remains confined in the tissues of the breast.

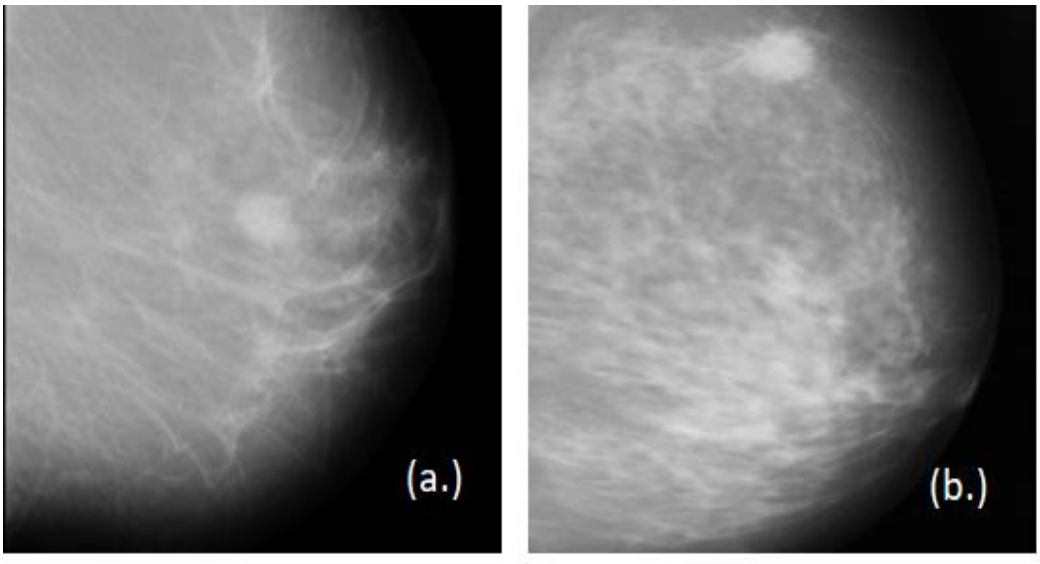

Fig.1.2: Categories of Tumor: (a.) Benign and (b.) Malignant

Invasive and non-invasive breast carcinomas are further categorized as follows[7]:

(a.) Ductal carcinoma in situ: In this category, cancer muscles have not been transmitted in the neighbouring tissue and are limited to the breast ducts.

(b.) Lobular carcinoma in situ: In this carcinoma also, cancer cells have not infected nearby tissues, but cells spread in the breast's milk-producing glands.

(c.) Invasive ductal carcinoma: It initiates in the tubes and can then propagate to neighbouring cells of the breast. This might extend to adjacent tissues in the body once it starts spreading beyond the breast ducts.

(d.) Invasive lobular carcinoma: This carcinoma starts in lobules and then infects adjacent tissues.

Some less frequent categories of breast cancer are as follows:

- **Phyllodes tumor:** This category of breast cancer is infrequent and mostly benign but can be malignant too. It can grow in the connected tissues of the breast.

- **Angiosarcoma:** Blood vessels or lymph vessels are the primary sources where this carcinoma grows in the breast.

- **Inflammatory breast carcinoma:** It is the rarest of a rare case and accounts for 1 to 5% of all cases. Having this carcinoma causes the breast to expand, feel boiling, and become red.

- **Triple-negative breast carcinoma:** This type of carcinoma occurs when a tumor does not have extra human epidermal growth factor receptor (HER2) proteins on its surface and lacks estrogen and progesterone receptors. Hormonal therapy for breast cancer is not sufficient, so this type of carcinoma is difficult to treat.

1.1.3 Evolution Stages

Breast Cancer evolves in stages with time. The stages of breast carcinoma can be separated depending on the tumor size and how much it has grown in the nearby tissues. The Cancer evolves in five main stages:

(i) Stage 0

(ii) Stage 1(In-situ)

(iii) Stage 2(Localized)

(iv) Stage 3(Regional)

(v) Stage 4(Distant)

Figure 1.3 shows the details of breast cancer categorization[8]:

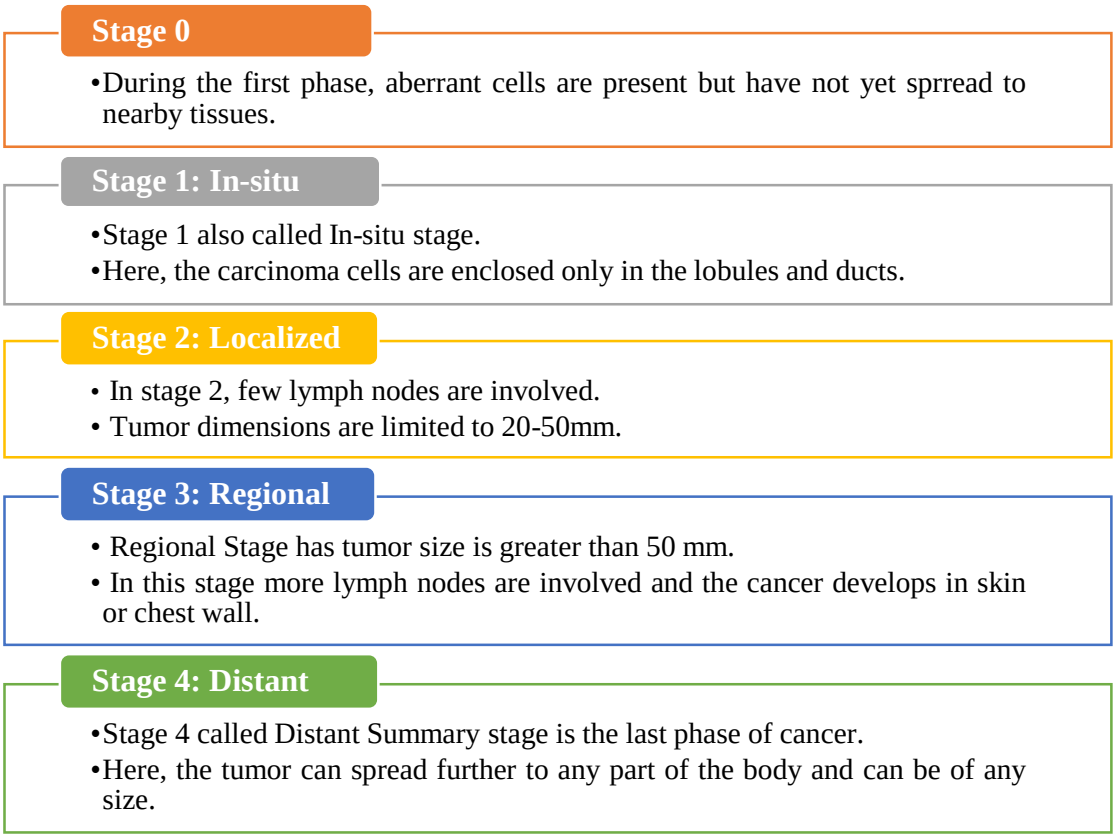

Fig.1.3: Evolution Stages of Breast Cancer

1.1.4 Breast Carcinoma Statistics

Mortality cases due to cancer are increasing at an alarming rate and are particularly affecting younger age groups. Compared to the 20 million cases in 2022, almost 35 million additional cancer cases are anticipated in 2050[9]. Figures 1.4 and 1.5 exhibit the cancer cases Worldwide and in India for the year 2022. As we can see, breast carcinoma is the most prevalent malignancy, accounting for 23.8% and 26.6% of all cases globally and in India, respectively[10].

A new Global Breast Cancer Initiative Framework has been suggested by World Health Organization, to save the lives of 2.5 million women from breast carcinoma by 2040[10]. A 2.5% yearly decline in breast carcinoma mortality will avert 25% of deaths by 2030 and 40% of fatalities by 2040 among females under the age of 70. To achieve these objectives, three essential elements are required: quick diagnosis, holistic management, and healthcare awareness for early detection.

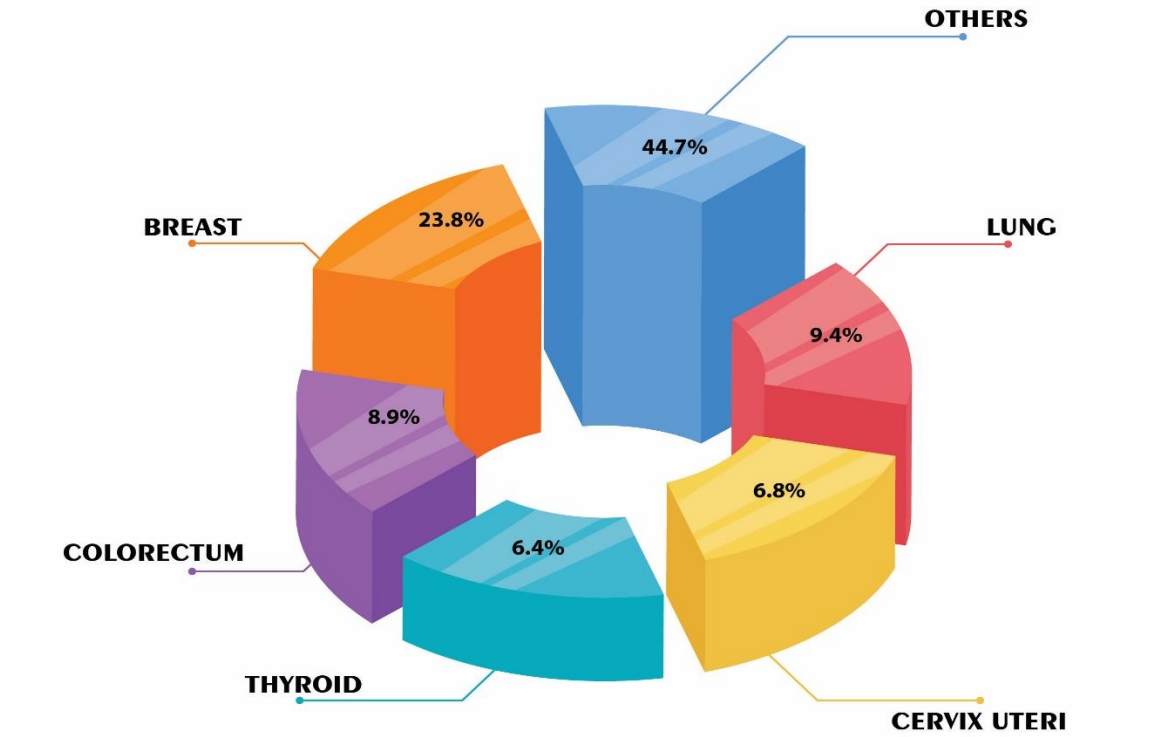

Fig.1.4: Cancer Statistics(World-Wide)-2022[10]

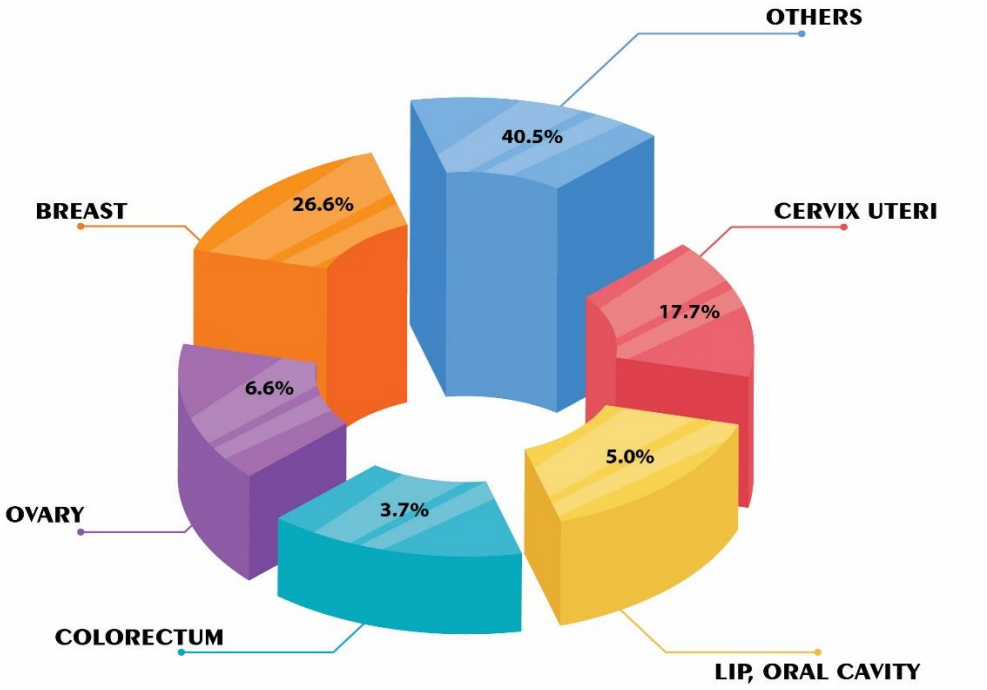

Fig.1.5: Cancer Statistics(India)-2022[10]

1.1.5 Breast Carcinoma Screening: Imaging Modalities

Breast Carcinoma screening could be done using various imaging modalities which are discussed as follows[11]–[13]:

(i) Screen Film Mammography(SFM)

SFM is the x-ray picture of the breast that is stored on a hard file on a phosphorous-coated film. Here, the contrast resolution is very poor and sensitivity in detecting cancer with dense breast is limited. This was considered the standard modality some years back, but now has replaced by digital mammography.

(ii) Full-field Digital Mammography/ Digital Mammography(DM)

In this imaging modality, digital detectors convert x-ray films into mammographic photographs of the breast (digital images). Mammography detects cancer at an early stage and is considered to be an effective and standard modality for early-stage breast cancer diagnosis by reducing the cancer rate by 15%.

(iii) Digital Breast Tomosynthesis(DBT)

It is also referred to as 3-D mammography. Breast images are acquired at multiple angles and thus create 2D and 3D pictures of the breasts. DBT detects cancer in women with dense breasts. But it is very expensive and the average radiation dose is 1 to 2 times higher than DM. Still, research is underway to determine whether 3-D mammography is better than 2-D mammography.

(iv) Ultrasound(US)

Ultrasound images are also called sonograms. It uses high-frequency sound waves and thus no radiation is involved. It can be used for pregnant women. It is the best way to figure out whether the anomaly is solid or fluid-filled. Ultrasound is still not considered as an early-stage screening modality as it may miss some solid mass during diagnosis.

(v) Magnetic Resonance Imaging(MRI)

MRI employs potent powerful magnetization and radio signals to generate high-quality pictures of the breast. MRI is mostly utilized for highly-risky individuals and it identifies incredulous areas that can further be used for biopsy. Secondly, it is very expensive when compared with ultrasound and digital mammography.

(vi) Histopathological Images

These are Hematoxylin & Eosin-stained images. In this, samples are taken from unusual breast regions (called a biopsy) and are observed under the microscope. This is used for diagnosing different breeds of cancer instead of detecting only malignancy because of multi-color images as it provides a comprehensive study of tissues. However high proficiency is required as the manual analysis is tedious and time-consuming.

1.2 Medical Image Processing using Computer-Aided Diagnosis(CAD)

Medical Images are very complex. Manual examination of medical images is quite difficult. Earlier, medical image processing was done manually. Still, it faces three main issues—firstly, lack of availability of multiple pathologists at one location. In addition, the process of manually analyzing images is arduous and unpleasant. Lastly, the diagnosis of breast carcinoma heavily depends on the pathologists' expertise and domain knowledge. Thus, to overcome misdiagnosis in the initial stages, Computer-Aided Diagnosis(CAD) acts as a "second notion" for solving breast cancer multi-classification problems[11]. Classifying medical images into distinct categories through CAD has become a useful tool for clinicians to aid in early diagnosis and therapy[14]–[16].

CAD systems require the processing of medical images; thus, extensive computational algorithms must be developed to process those images. CAD is reasonably good at detecting invasive breast cancers in their early stages but increases the peril of false positive results. Figure 1.6 depicts the CAD as an interface or a mediator connecting the medical science branch with computer science. Thus, accuracy and workflow efficiency must be considered before implementing CAD in clinical practice, and secondly, to avoid improper use, user education is crucial to grasp the features and constraints of CAD systems [17]. There are two types of CAD systems [18]:

(i) **Computer-aided detection (CADe) systems:** These systems provide automated anomaly identification and localization in medical imaging data.

(ii) **Computer-aided diagnosis (CADx) systems:** These frameworks are utilized as an extension of the detection model, offering more data and facilitating decision-making. They can be used to categorize whether an abnormality is benign or malignant.

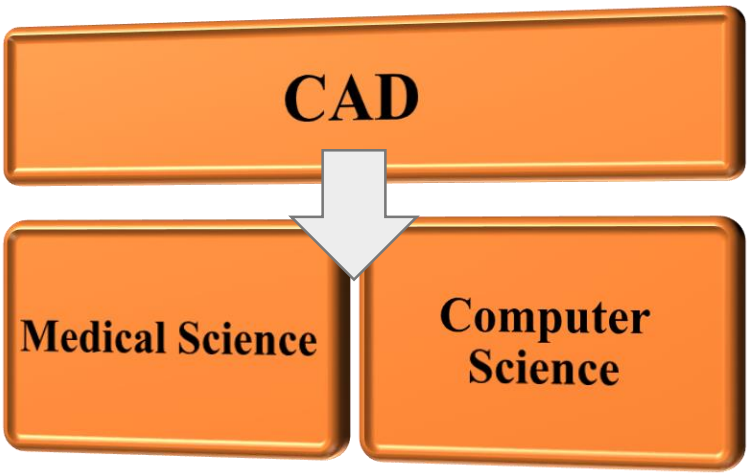

Fig.1.6: CAD serves as a bridge between computer science and medicine.

1.3 Machine Learning & Deep Learning in Medical Image Processing

In the previous few decades, there has been significant growth in creating sophisticated algorithms and effective preprocessing approaches in Machine Learning (ML) and Deep Learning(DL). Among these improvements is the evolution of artificial neural networks (ANNs) into more sophisticated designs known as deep neural networks (DNNs)[19].

In this context, ML and DL have created methods that can diagnose the illness more precisely at the initial stages thus diminishing the number of readmissions in hospitals and clinics. Deep learning tackles a broad range of issues in healthcare, such as personalized therapy recommendations, infection monitoring, and cancer detection[20]. Thus, the adoption of artificial intelligence (AI) tools can facilitate the acquisition of new fidelity procedures and lower the expense of healthcare resulting from inaccurate diagnoses [21].

In the field of medical imaging, DL has achieved tremendous progress, attaining remarkable outcomes in several tasks. There is still an obstacle in the form of the restricted availability of training information, especially in the healthcare domain where obtaining data can be expensive and governed by privacy laws[22]. Image mining, computer vision, and pattern recognition have all become more important aspects of medical image processing[20].

1.3.1. Machine Learning-Based CAD System

CAD system implementing machine learning comprises the following steps[23], [24]:

(i) Image pre-processing: This step includes eradicating the noise, clearing the picture, and getting it ready for the next step

(ii) Image segmentation: It includes partitioning the picture into several parts, and focusing only on the specific regions called regions of interest (ROI).

(iii) Feature extraction and selection: It comprises identifying and retrieving valuable characteristics from the previously processed pictures. Among the many benefits of feature selection include lower costs, shorter training periods, and higher accuracy.

(iv) Classification: This step assigns labels or classes to distinct groups through the use of different classifiers

1.3.2. Deep Learning-Based CAD System

DL falls within the category of the representation learning approach[25]. With multiple non-linear processing layers, deep learning extracts features directly from the data[15]. Contemporary research communicated that "Convolutional Neural Network (CNN)" attains remarkable performance in cancer detection and diagnosis[26]. Layers adjoining the input layer learn low-level features and are more generic. In contrast, layers adjoining the output layer learn distinctive features of the input image and, thus, are more specific[27].

Steps followed for DL based CAD system are discussed below[23]:

(i) Image pre-processing: This is the same step followed in ML-based CAD systems i.e. processing the image for noise removal and preparing it for the subsequent stage.

(ii) Convolutional layers: This is also called the feature extraction step. It includes convolution operations to instinctively extract high and low-level characteristics from the image.

(iii) Pooling layers: These are added to reduce the spatial dimensionality of attributes.

(iv) Fully connected and Softmax layer: These layers perform categorization in CNN.

The specific aspects of these layers are covered in Chapter 3. When inadequate data is available to train a generalizable framework, data expansion is executed to enlarge the statistics of data points. Using artificial synthesis, data augmentation can provide more samples, thereby expanding the training set. In situations when rich training sets are not available, this deeply ingrained method in computer vision has proven indispensable[28]. Operations performed for augmentation involve rotation, noise inclusion, image cropping, and other geometrical activities[29]. Figure 1.7 shows the ML and DL-based CAD systems.

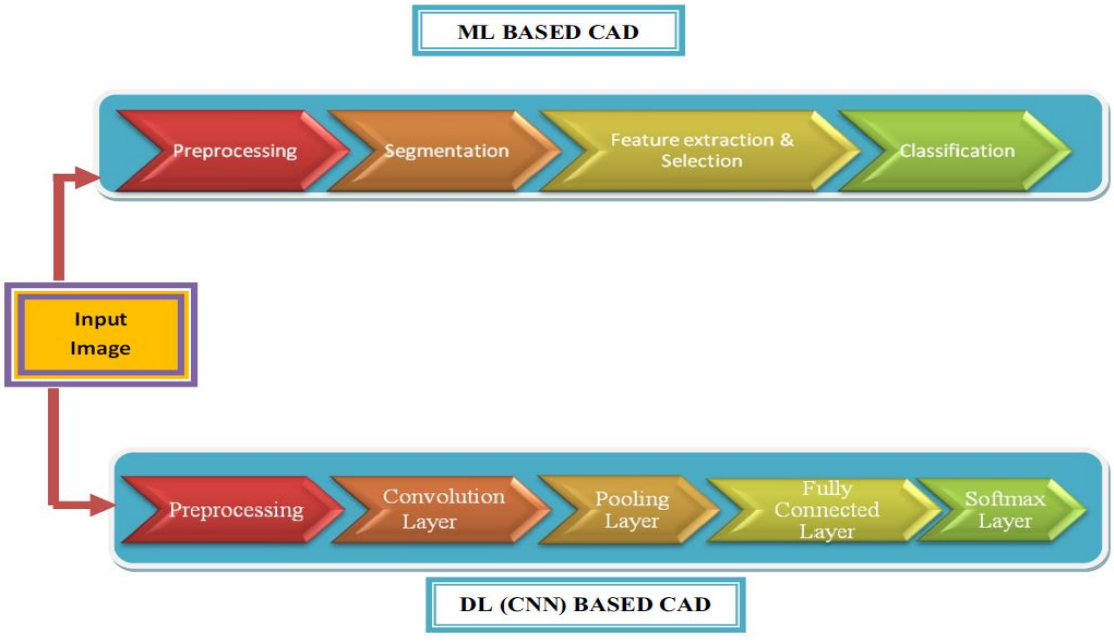

Fig.1.7: ML and DL based CAD system

1.4 Research Gaps & Challenges

Globally, breast carcinoma is climbing at a frightening pace. In contrast to developed countries, fatality estimates are comparably high in low-wage and middle-wage countries[5], [30]. The significance of early diagnosis and detection is critical in improving long-term survivability[11]. The following research gaps are identified while doing the literature survey.

The challenges and gaps identified in the literature are discussed as follows:

(i) Accuracy of Deep Neural Networks on Medical Datasets

In the literature, research is performed for diagnosing malignancy using ML and DL strategies. Still, the researchers cannot attain a remarkable performance with ML and DL models. There is a lack of Models trained on mammography images with Cranio-Caudal (CC) and Medio Lateral Oblique(MLO) with good accuracy values.

Thus a need arises to develop a competent architecture with deep neural networks to attain remarkable performance for diagnosing & classifying breast cancer using ML and Neural Net classifiers.

(ii) Utilization of smaller datasets in the medical domain

Deep Neural Networks require abundant data during training phases. Due to this, training DNNs with smaller medical datasets is still a challenge. Transfer learning has been deployed in many research studies. However, training deep models in multiple stages on smaller datasets has not yet been explored.

Thus, a need arises to develop a framework to train DNNs on smaller and different datasets with multiple stages and evaluate their performance.

(iii) Generalizability of Deep Neural Networks(DNNs)

DNNs can memorize the training data and are prone to overfitting. Generalizability defines the execution of the model on unknown data. Very less work has been focused on the generalizability & generalization error of DNNs. However, analyzing the exact generalization error is practically impossible as models behave separately in every scenario.

Still, there is a need to approximate these errors and thus develop an approach for assessing the generalizability of DNN on previously unknown data.

1.5 Motivation for Study

As we discussed the Breast Cancer Statistics in section 1.1.3, we observed that breast carcinoma is the most prevalent malignancy in the country which has surpassed cervical cancer. Figure 1.8 reveals the startling rise of breast carcinoma. In India, the late diagnosis of the disease results in low survival rates [30], [31]. The people usually don't go for screening until any symptom appears. Age-specific breast cancer occurrences are shown in Figure 1.9. Younger age groups are more impacted by this fatal disease, as we can see from the comparison between the situation 25 years ago and the present. Research Gaps discussed in Section 1.4 ascertain that researchers are still competing to attain remarkable performance using ML & DL. Thus, there is an urgent need to propose robust models that attain outstanding performance and thus could be deployed for early-stage diagnosis. These models could assist doctors and help patients who are diagnosed with the disease and are looking for a second opinion. Thus, could be useful in providing timely diagnosis thus leading to increased survivability in patients.

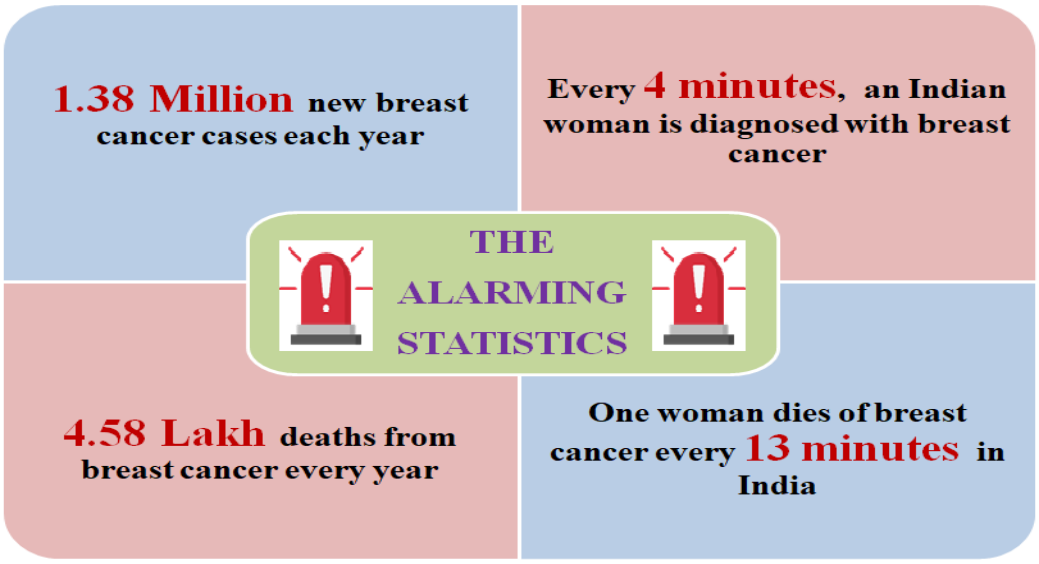

Fig.1.8: Alarming Statistics

(Source: https://cancerconsultindia.com/blog/breast-cancer-statistics-rise-of-breast-cancer-in-india)

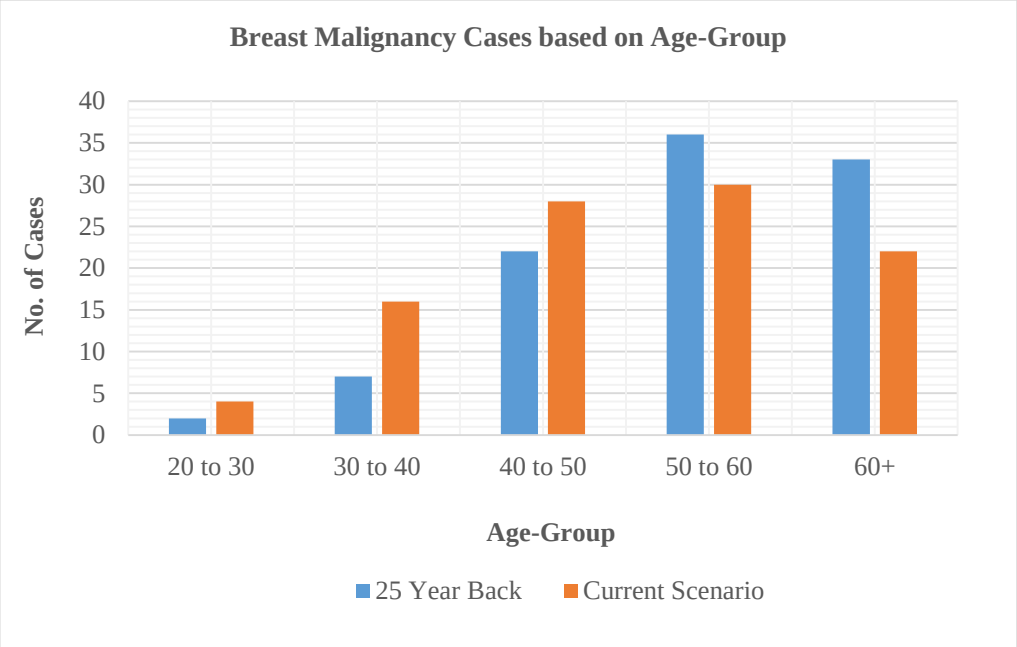

Fig.1.9: Incidences of Breast Cancer as per age-group in India[32]

1.6 Problem Statement

Breast Carcinoma is increasing at an exponential pace. Section 1.1.4 and 1.5 illustrate the alarming statistics. As compared to 20 million cases in 2022, there is an anticipation of 77% rise of breast malignancy cases by 2050[9]. Research Gaps discussed in section 1.4 depict that there is an urgent need to propose robust models that attain commendable performance and thus could be deployed for early-stage diagnosis. Additionally, there is a need to propose efficient frameworks that utilize smaller datasets in the medical domain and attain remarkable outcomes. These CAD systems could assist doctors and help patients who are diagnosed with the disease and are looking for a second opinion. Further, there is necessity to evaluate the generalizability of deep neural networks.

Thus, the problem statement for this study can be stated as: Developing efficient models for early stage breast cancer diagnosis using deep learning strategies and evaluating generalization error in deep neural networks to assess their generalizability to unseen data.

1.7 Research Objectives

We aim to propose a framework for the diagnosis and classification of breast carcinoma using DL. Following are the objectives of the research that fulfills the stated aim:

(i) To conduct a Systematic Literature survey on the diagnosis & classification of Breast Cancer using existing Machine Learning and Deep Learning approaches.

(ii) To propose a model for classifying Breast Cancer using pre-trained networks and examine various evaluation parameters.

(iii) To propose a framework for implementing Multi-Stage Transfer Learning for classifying Breast Cancer and evaluate the results.

(iv) To develop an approach for the estimation of generalization error for Deep Convolutional Neural Networks.

(v) Comparison of the findings obtained from the proposed methodology to other current approaches

1.8 Organization of Thesis

The dissertation comprises six chapters discussing the research work in a concise & understandable way. A summary for each chapter is outlined below:

Chapter 1 presents an overview to Breast cancer and Medical Image Processing using computer-aided diagnosis (CAD). This chapter throws light on breast carcinoma categorization, stages, and screening modalities. It discusses the significance of early-stage tumor diagnosis using ML and DL strategies. Research Gaps and Research Objectives are also mentioned.

Chapter 2 describes the survey in the domain of Breast Carcinoma Diagnosis using ML and DL. Instances of breast malignancy are rising at a frightening rate. Mortality count could be reduced if diagnosis takes place on time. Extensive research has been covered in the literature for diagnosing breast malignancy in the early stages. But still, there is a gap to boost the performance of early-stage diagnosis through DL strategies.

This chapter explores the strategies adopted by the researchers along with their merits and de-merits for diagnosing & classifying breast carcinoma using several ML and DL strategies. It also describes the datasets and deep neural networks available for training on medical image datasets.

The following paper has been published from this work:

- G. Chugh, S. Kumar, and N. Singh, "Survey on Machine Learning and Deep Learning Applications in Breast Cancer Diagnosis," Cognitive Computation. 2021, doi: 10.1007/s12559-020-09813-6. **(SCIE Indexed, IF=4.3)**

Chapter 3 proposes a dual framework called Transnet for diagnosing breast carcinoma using Machine Learning and Deep Learning paradigms. Two separate experiments were performed. The research was carried out on the CBIS-DDSM dataset. In the first approach, i.e., Deep feature extraction with ML classifier, Deep Convolutional Neural Network(DCNN) models like VGG-16, VGG-19, ResNet-50, and ResNet-152 are deployed as feature extractors, and the obtained features are utilized for training conventional machine learning classifiers. The second approach, called Deep Learning feature extraction with a neural network classifier, exploits MobileNet, VGG-16, VGG-19, ResNet-50, ResNet-152, and, DenseNet-169 for feature extraction and categorization. The merits and demerits of both approaches are discussed along with the futuristic suggestions. The chapter contrasts the stated framework with other cutting-edge strategies and highlights the significance of the proposed strategy.

The following paper has been published from this work:

- G. Chugh, S. Kumar, and N. Singh, "TransNet: a comparative study on breast carcinoma diagnosis with classical machine learning and transfer learning paradigm," *Multimed. Tools Appl.*, no. 0123456789, 2023, doi: 10.1007/s11042-023-16938-x. **(SCIE Indexed, IF=3.0)**

Chapter 4 illustrates another proposed architecture for the early diagnosis and categorization of breast malignancy using the Multi-Stage Transfer Learning Approach(MSTLA). Utilizing smaller datasets and achieving remarkable performance is still a challenge. To achieve this, we are utilizing the concept of transfer learning with multiple stages. This approach is generally followed when the data required for training a model is unavailable in abundance. Experiments were carried out on three mammography datasets- Mammography Image Analysis Society (MIAS), In-Breast, and CBIS-DDSM. Results demonstrate the deployment of the framework in real-life scenarios. The chapter portrays the analysis of the stated framework with other cutting-edge approaches.

The following paper has been published from this work:

- G. Chugh, S. Kumar, and N. Singh, "MSTLA: Multi-Stage Transfer Learning Approach for Breast Carcinoma Diagnosis," *2023 Int. Conf. Adv. Comput. Comput. Technol. InCACCT 2023*, pp. 509–514, 2023, doi: 10.1109/InCACCT57535.2023.10141697.

Chapter 5 focuses on the generalizability and generalization error in deep neural networks. Generalization is an approach to analyze how the model behaves on unseen data. Generalization Error(GE) measures the difference between training and testing errors. Leading causes of GE include memorization of training data, overfitting, a model with too many parameters, etc. We have presented a model for assessing the generalizability of Deep Convolutional Neural Networks through various noises. Gaussian, Salt and pepper, and Speckle noise were added to the CBIS-DDSM dataset. Generalizability and Generalization Error were evaluated for three DCNNs - Inception Net v3, DenseNet-201, and EfficientNet B4.

The following paper has been published from this work:

- G. Chugh, S. Kumar and N. Singh, "A Framework for Generalization Error Evaluation in Deep Convolutional Neural Networks," *2023 IEEE Engineering Informatics*, Melbourne, Australia, 2023, pp. 1-7, doi:

10.1109/IEEECONF58110.2023.10520375.

Chapter 6 concludes the research findings along with prospects & social impact. The chapter specifies the impact of the proposed research work on society and throws light on the application areas for future researchers in this field.

1.9 Chapter Summary

The chapter discusses the basic concepts of breast cancer along with the categorization and evolution stages. Breast Cancer could be screened using various imaging modalities. Statistics have shown that breast malignancy is the primary reason of demise in females. Thus, an appropriate strategy needs to be developed to diagnose the deadly disease at an initial stage. CAD using ML and DL is an hour of need. The chapter also discusses the research gaps followed by research objectives formulated to fill the gap. The chapter concludes with the thesis organization throwing light on the summary of the chapters drafted in the dissertation.

CHAPTER 2

METHODOLOGICAL LITERATURE SURVEY

2.1 Overview

This Chapter covers the systematic literature survey in the field of breast carcinoma diagnosis. Section 2.2 highlights the strategies for utilizing Deep Convolutional Neural Networks. Survey Analysis is presented in Section 2.3. Sections 2.4 and 2.5 discuss the related work in the domain using ML and DL strategies. Generalization approaches are elaborated in Section 2.5.

2.2 Deep Convolutional Neural Networks(DCNN)

2.2.1 Introduction

Artificial Intelligence(AI) has emerged as the most promising field for various types of research in the current industries. Deep Learning and Machine learning, the subfields of AI, are giving tremendous results in each & every sector. We also use these applications in our daily lives, like scrolling the search engines, taking to digital assistants, playing innovative games, and using social media apps. Etc. In recent years DL and ML are also been widely used in the medical sector. These advanced technologies are helping doctors in the treatment, reducing the diagnosis time and thus saving the lives of patients.

In Cancer Detection and Diagnosis, Convolutional Neural Network(CNN) has performed remarkably. These networks consist of multilayer neurons capable of recognizing valuable features and thus aiding detection and classification. ‘Deep CNN i.e. DCNN’ refers to the layers in the network[33]. Initial layers learn generalized

characteristics from the pictures, and the deeper layers learn more specific attributes. A generalized model for CNN is illustrated in Figure 2.1.

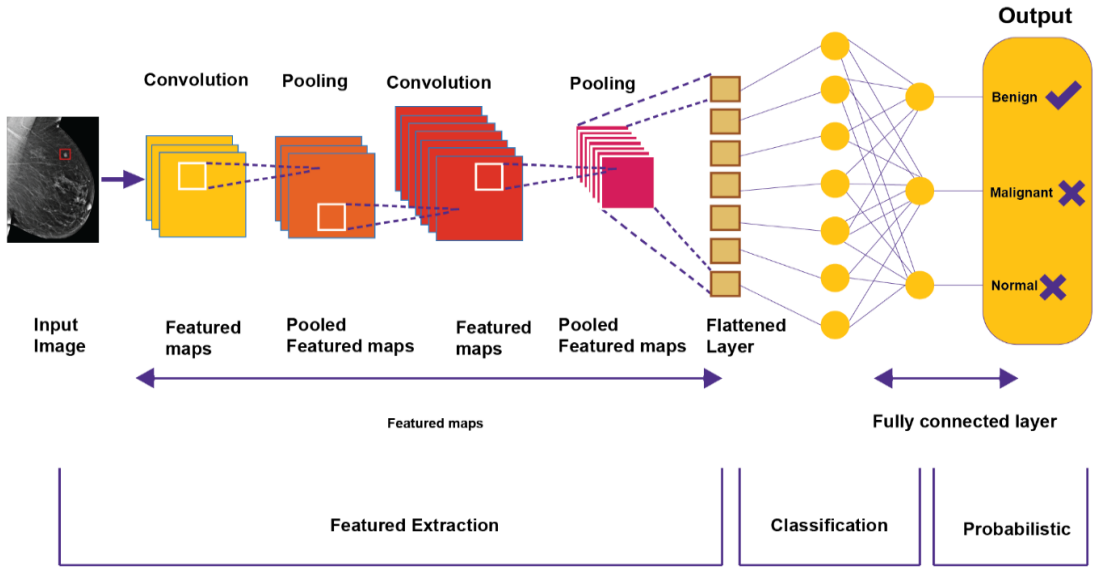

Fig. 2.1:CNN Architecture[34]

(Source: https://www.analyticsvidhya.com/blog/2022/01/convolutional-neural-network-an-overview/)

Layers used in CNN are discussed as follows[35], [36]:

(a.) Convolutional Layer: The convolutional layer holds primary importance in CNN design. It consists of several convolutional filters, sometimes referred to as kernels. The convolution step yields the essential characteristics that are derived from the pictures.

(b.) Pooling Layer: These layers lower the spatial dimensions of the image by down-sampling the convolved features. This approach reduces the size of large-scale feature maps to produce smaller feature maps. Thus it accomplishes the two goals i.e. enhancing the feature abstraction and lowering the computational capacity required for processing the information. Two popular poling techniques are Average Pooling and Max Pooling. The result of average pooling is the aggregate of all the data points from the image area the kernel

covers. In contrast, in Max-Pooling, a pixel's maximum value is chosen from a region of the picture that the kernel covers.

(c.) Activation function Layer: Layers of non-linear activation are applied after all weighted layers. These layers facilitate non- linear input to output mapping and thus helps the CNN to learn more challenging tasks. Several activation functions are used in CNN which as discussed as follows:

- Sigmoid: Upon receiving real numbers as input, this activation function only produces values ranging from 0 to 1. The sigmoid function curve can be expressed mathematically as:

$$\text{sig}(x) = \frac{1}{1+e^{-x}} \tag{2.1}$$

- Tanh: This activation's output is limited to −1 and 1. It is represented as:

$$\tan(x)=\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \tag{2.2}$$

- ReLU: This is most frequently utilized within the CNN as it is computationally efficient. It changes all of the supplied values to positive integers. ReLU's primary advantage over the others is its lower computational burden. The following is an illustration of its mathematical representation:

$$\text{rel}(x)=\begin{cases}0, & x < 0\\ x, & x \geq 0\end{cases} \tag{2.3}$$

- Softmax: To normalize a neural network's output to a probability distribution over anticipated output classes, the Softmax function is frequently employed as the final activation function. It translates raw output values (logits) into probabilities. It is represented as:

$$\text{soft}(\overline{p_i}) = \frac{e^{p_i}}{\sum_{j=1}^{k} e^{p_j}} \tag{2.4}$$

Here, $\overline{p_i}$ – input vector

e^{p_i} –exponential function for an input vector

e^{p_j} - exponential function for output vector

k- number of classes

(d.) Batch Normalization Layer: Convolutional neural networks frequently employ the Batch Normalization (BN) layer to normalize each neuron's input so that its mean and variance are both zero. This aids in stabilizing the learning process and guards against internal covariate shift (A problem arises when a layer's input distribution varies because of weight updation during training). Since it has little effect on regularisation, it avoids the vanishing gradient problem and lowers the chance of over-fitting.

(e.) Classification Layer: The main objective is to perform classification and produce the final class as output. Also named the Fully Connected(FC) layer which links every neuron to all neurons in the preceding layer. After the flattening operation, FC layers receive a vector as input from the final pooling or convolutional layer in the network.

2.2.2 Methods to Implement Deep Convolutional Neural Networks

DCNNs can be implemented through the following two mechanisms:

(i) Training from scratch: When training a CNN in this approach, a significant amount of training sample is fed to the network so that the network can understand the attributes right from the beginning. This process requires selecting suitable layers, hyper-parameters, optimizers, etc. It is a time-consuming process and involves processing on powerful GPUs.

(ii) Transfer Learning: This strategy of training a CNN is used when there are fewer training instances in the target class. Various pre-trained networks could be utilized for training the model. These networks are already trained on massive datasets; thus, the network has learned the generic features. Therefore, the knowledge learned from the base domain is transferred or used for training the destination domain where training samples are less. As a result, this approach requires less time and thus could be used on CPUs.

Figure 2.2 visualize CNN Implementation Techniques.

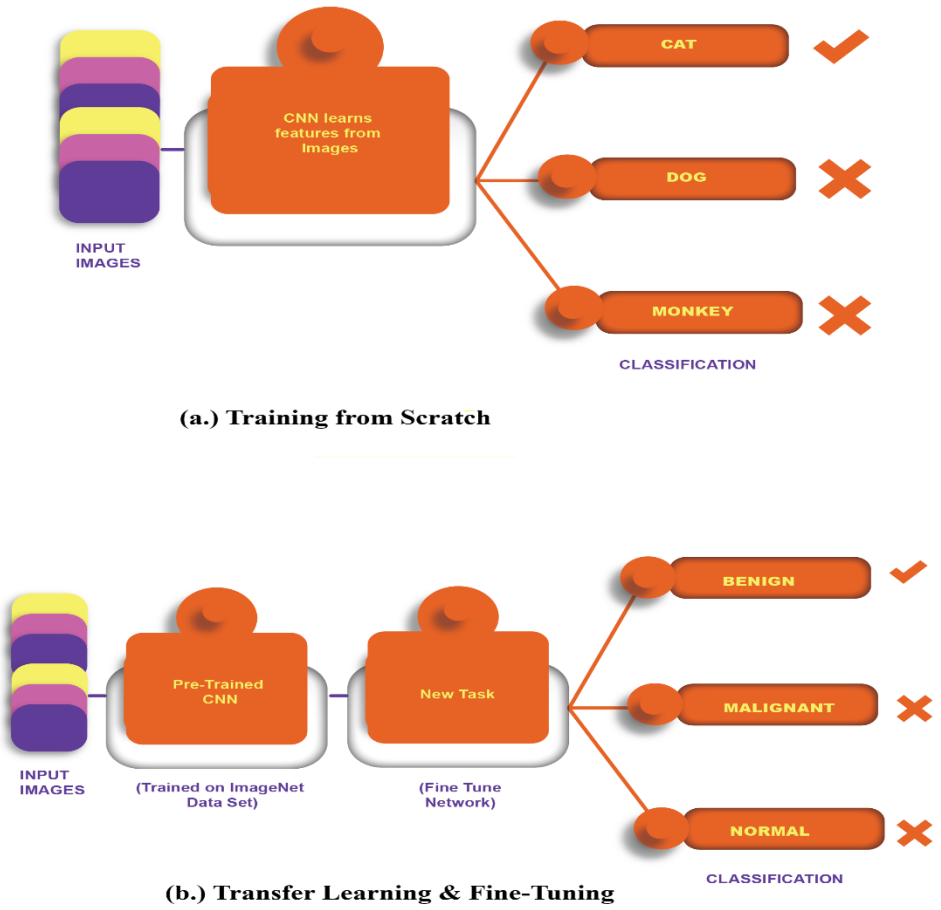

Fig. 2.2: CNN Implementation[37]

2.2.3 Strategies to Utilize Pre-Trained DCNN

Section 2.2.2 discusses the various methods to utilize DCNNs. The Transfer Learning approach can be further implemented in the following three ways [13]:

(i) **Baseline Model:** In this category, the complete model is trained from beginning to end, and only the structure of the pre-trained network is exploited (Fig 2.3(a)).

(ii) **Fine-tuning:** This includes transferring weights from a pre-trained network to the destined model and could be accomplished in two methods: Layer by Layer and Partial fine-tuning of the model. Layer-level tuning initiates with the outermost layer. Then additional layers are trained in chronological order whereas, in partial training, the weights of the initial layers are left unchanged, and the upper layer's weights are modified to train the unfamiliar dataset (Fig 2.3(b)).

(iii)**Feature extractor:** This approach utilizes a pre-trained network's convolutional base in its original form, with no changes to its specified weights. Traditional classifiers substitute the dense layers of the pre-trained network. The convolutional base outcomes are passed directly to the train classifiers. (Fig 2.3(c))

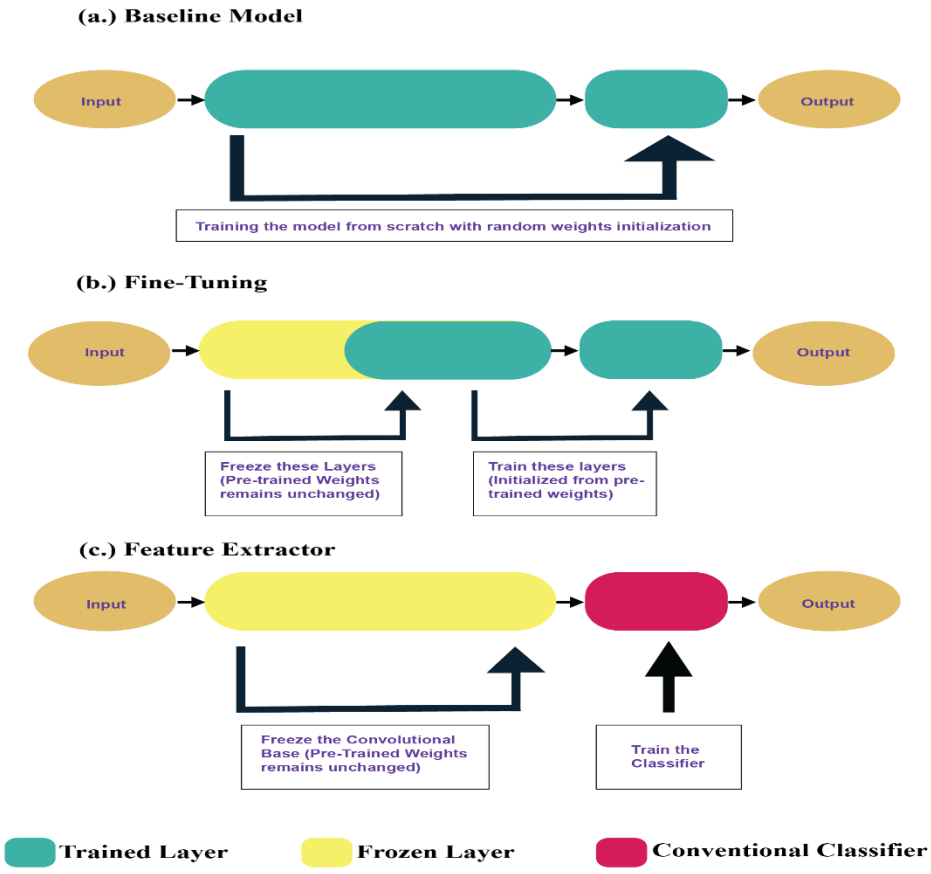

Fig. 2.3: Strategies for Transfer Learning Implementation[38]

2.2.4 Performance Measures

Various measures have been defined to assess the efficiency of any ML/DL system. The Confusion matrix indicates the relationship between predicted and actual values (Table 2.1).

Table 2.1: Confusion Matrix

	Predicted Positive	**Predicted Negative**
Actual Positive	True Positive	False Negative
Actual Negative	False Positive	True Negative

In the above table,
True Positive (TP) – a person diagnosed as positive and is genuinely positive.
True Negative (TN)- a person diagnosed negative and is actually negative.
False-positive (FP)- a person diagnosed positive but is truly negative.
False Negative (FN)- a person diagnosed as negative but is genuinely positive.

Following is an overview of the measures [13]:

(i) Accuracy: It is the proportion of correctly categorized instances to the overall samples and is interpreted as:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad (2.5)$$

(ii) Sensitivity/Recall/True Positive Rate (TPR): It evaluates what percentage of patients are correctly identified to have a particular disease (+ve instances) and illustrated as:

$$\text{Sensitivity (Recall)} = \frac{TP}{TP+FN} \qquad (2.6)$$

(iii) Specificity/True Negative Rate (TNR): Determines what proportion of persons are correctly identified not to have a particular illness (-ve instances) and elucidated as:

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (2.7)$$

(iv) Precision/Positive Predicted Value (PPV): Specifies what percentage of patients are actually relevant(+ve) and described as:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2.8)$$

(v) F-Measure: Interpreted as the harmonic average of precision and recall and is illustrated as:

$$\text{F-Measure} = \frac{2*\text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}} \quad (2.9)$$

(vi) False Positive Rate (FPR): It anticipates a favourable outcome when the finding is actually negative. FPR is given as:

$$\text{FPR} = \frac{FP}{TN+FP} \quad (2.10)$$

(vii) Area Under the Receiver Operating Characteristic Curve (AUC-ROC): It defines a network's capability to distinguish between different categories. It is drawn by coordinating TPR on the y-axis and FPR on the x-axis. Thus, the network's performance is examined by determining the area under this curve, whose value lies from 0 to 1. The higher the value, the more powerful and reliable the network is at distinguishing positive individuals as positive and negative individuals as negative.

2.2.5 Loss Function

One of the most crucial components of neural networks is the loss function, which works directly with optimization functions to adapt the network to the provided training information. The efficiency of a DNN can be assessed by its loss function. The loss function in DL tasks typically quantifies the degree of precision, or fit between the ground-truth value and the predicted value. The goal of training is to reduce this loss between the desired outcome and anticipated outputs. To minimize the average loss, the hyper-parameters are accordingly adjusted[39].

For classification problems two loss functions are used which are discussed as follows[39]:

(a.) Binary-Cross Entropy(BCE)/ Log Loss

This function is applied in two-class categorization scenarios, where the model must classify the input into one of two predetermined categories. The loss function is given as:

$$\text{BCE Loss} = \frac{1}{n}\sum_{i=1}^{n} - [y_i \,.\, log(p_i) + (1 - y_i) \,.\, log(1 - p_i)] \qquad (2.11)$$

Here, n is the count of data points, y_i signifies the actual category value (i.e. 0 or 1) and p_i depicts the Softmax probability of any instance in that particular class (i.e. 0 or 1).

(b.) Categorical-Cross Entropy(CCE) Loss

This loss is utilized in multi-class categorizations. The functionality is similar to binary cross-entropy loss and is stated as:

$$\text{CCE Loss} = -\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{c}[[y_{i,j} \,.\, \log(p_{i,j})] \,] \qquad (2.12)$$

In the above equation, n is the count of data points, c represents the number of classes, y_i signifies the actual category values and p_i depicts the predicted probability of any instance in that particular class. Thus, in CCE Loss evaluation, the loss for each class is evaluated independently and then combined to get the overall loss.

. 2.3 Survey Analysis

This section discusses the papers reviewed and the electronic repositories explored during the literature survey. It also covers the database repositories utilized by the researchers.

2.3.1 Papers Explored/Reviewed

Breast Cancer particularly affects younger people, and there is a crucial need to diagnose it early. In this research, more than 288 papers were peer-reviewed. 50-60 articles diagnose breast carcinoma using machine learning classification algorithms while the rest implement deep learning approaches. Figure 2.4 shows the statistical distribution of ML and DL articles referred to in this research.

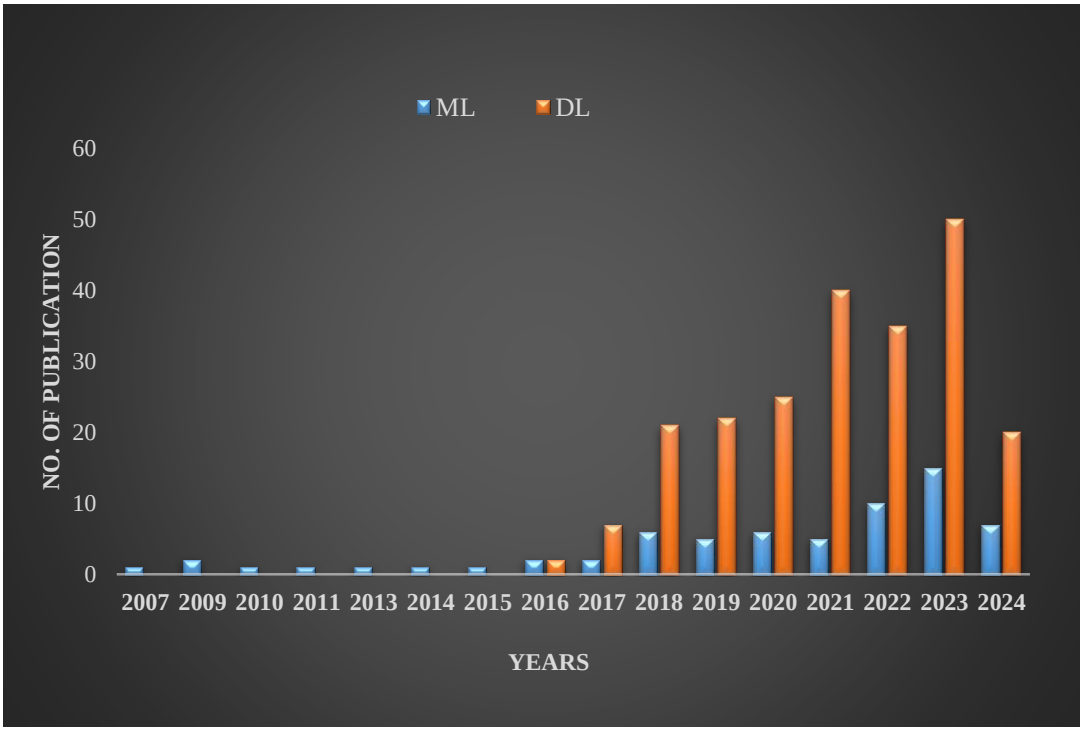

Fig. 2.4: ML and DL publications examined in the survey

2.3.2 Electronic Repositories

This research analyses several studies on ML and DL strategies to diagnose and classify breast carcinoma. Searching criteria for inclusion and exclusion of the peer-reviewed journal paper is based on the following keywords: breast carcinoma, computer-aided diagnosis, machine learning, deep learning, classification, detection, computer-assisted diagnosis, and computer-aided detection. The researchers' electronic databases used in the survey include Elsevier, Springer, IEEE, Wiley,

Oxford, PLOS One, Hindawi, Taylor & Francis etc. The articles not meeting the required criteria were excluded and the rest were considered during the survey.

2.3.3 Database Repositories

The datasets utilized by the researchers in the survey are illustrated in Table 2.2.

Table 2.2: Database repositories

Digital Repository	Dataset Name	Dataset Description
UCIMLR/ University of California Irvine Machine Learning Repository[40]	WBCD/Wisconsin Breast Cancer Dataset	WBCD is a widely used dataset among experts working on machine learning research. It contains 569 instances with 32 attributes
BCDR/Breast Cancer Digital Repository[41]	a.) BCDR-FM/ Film Based Digital Repository b.) BCDR–DM/ Full Field Digital Mammography	It comprises mammography and ultrasound images of 1734 patients (1010 –BCDR-FM and 724 – BCDR-DM) along with their clinical history.
MIAS / Mammography Image Analysis Society[42]	MIAS / Mammography Image Analysis Society	It has 322 digitized mammography breast images at 50-micron resolution
The University of South Florida[43]	DDSM / Digital Database for Screening Mammography	It comprises 2620 cases which are annotated in 3 categories- normal, benign, malignant
BreakHis[44]	BreakHis / Breast Cancer Histopathological Database	It comprises 7909 microscopic pictures. 7909 cases are divided into 2480 benign samples and 5429 malignant samples

SEER[45]	SEER / Surveillance Epidemiology and End Result	SEER dataset provides stage-wise survivability statistics of breast cancer data
InBreast[46]	InBreast	This database consists of 115 cases with 410 images.
IRMA[47]	IRMA/ Image Retrieval in Medical Applications	It consists of 10,509 reference breast images

2.4 Methodological Survey on Breast Tumor Diagnosis Using Machine Learning Classifiers

This section covers an analysis of research conducted by multiple authors using different machine-learning classifiers to diagnose breast carcinoma [13].

Gene Microarray datasets were used by Osareh and Shadgar[48] to classify benign and malignant lesions. The researchers extracted attributes using Principal Component Analysis (PCA), ranked features using Signal Noise Ratio (SNR), and feature selection was implemented using a wrapper-based approach, i.e., Sequential Forward Selection (SFS). The authors concluded that the best results could be obtained by appropriate selection of kernels and classifiers. SVM with the RBF kernel attains an accuracy of 96.33%. Their work suggested that SVM models with Radial Basis Function (RBF) kernel could be further explored as a diagnostic aid for breast carcinoma diagnosis.

Feature Selection and Reduction are very crucial parameters that affect classifier performance. The finest classifier will perform worse if features are not chosen wisely. Karabatak and Ince [49] used Association rules (AR) on WBCD and reduced feature dimensions from nine to four. The authors obtained 95.6% and 97.4% accuracy with 4 and 8 inputs by performing classification using Neural Network (NN). K. Goyal et al. [50] implemented seven different algorithms - Bagging, Ada Boost, LR, SVM, RF, NB, and DT on WBCD. The authors compared classifiers' results by implementing various feature selection techniques. Ada Boost and LR classifier attained significant performance using PCA with an accuracy of 97.92% for the reduced dataset.

Similarly, [51] utilized WBCD and BCDR to evaluate four classifiers: NB, KNN, J48, and SVM. The Weka tool's "WrapperSubsetEval" wrapper feature selection method was employed on both datasets. Findings showed that the superior results were given by SVM with an accuracy of 97.91% and 95.10% on WBCD and BCDR. Some of the well-known research on breast carcinoma diagnosis is summarized in Table 2.3.

In certain situations, an ensemble model is created by merging multiple classifiers with various techniques to improve the results from individual classifiers. The researchers in [52] presented an ensemble approach with 97.8% accuracy that used SVM, DT, and LR as classifiers. Weights were assigned to each classifier using the Sequential Least Square Programming (SLSQP) technique. The outcomes show that the ensemble framework outperforms a standalone ML classifier. Kashif et al. [53] implemented several ML models -SVM, Ada Boost, DT, LR, RF, Gradient Boost, and KNN on the MIAS dataset. Results proved that SVM performs best with an accuracy of 90%. In contrast, the Gradient Boost classifier gave the worst efficiency with an accuracy of 52%. Another work on the MIAS dataset was proposed in [54], where authors applied discrete wavelet transform (DWT), PCA & SVM approach and attained an accuracy of 82.85+-2.21%. The proposed work could be continued in the future by applying other variants of wavelength transform, such as stationary wavelet or wavelet package transform. The authors also suggested using deep learning methods on large datasets in the future to obtain high accuracy and precision for prediction and classification.

Cancer prognosis refers to the probability of success of treatment leading to recovery. A lot of work on cancer prognosis has been conducted in the past few years. Similar work was proposed in [55] by evaluating eight different classifiers, firstly, by comparing their AUC value for the prognosis time of 2, 5, 8, and 11 years, and secondly, based on the importance of different variables. The authors concluded that, on average, a 5-year prognosis period shows the overall best performance. Out of eight models, overall RF, Boosted Trees, partial least square, and generalized linear model (GLM Net) classifiers provided the best results with an AUC value of approximately 0.75. In contrast, the KNN model performed worst with an AUC value of 0.67. In the future, their work could be extended by including bagging, boosting, and other ensemble techniques. The experts in [56] used data-driven knowledge for predicting 5-year breast cancer survival on the SEER database on all the BC instances in 1988–

2009. The authors compared stage-specific (in situ, localized, regional, distant) and joint predictive models for three distinct algorithms—NB, LR, and DT. The authors concluded that summary stage predictive models achieve superior results compared with the joint model; secondly, it is essential to collaborate all the information attained from various stages with its validity time.

Many times, oncologists face difficulties in the identification of the disease through the naked eye. Under such circumstances, the field of medicine called radiomics is utilized to examine radiological pictures and derive radiomic features, which reveal health abnormalities. Using machine learning and radiomic characteristics, the experts [57] categorized tumors on DBT pictures. Seventy radiomic characteristics were retrieved. The authors selected six features using the least absolute shrinkage and a selection operator (LASSO) because of lesser training data. In the case of 70 radiomic features, SVM shows the best results with a 55% recognition rate for a non-cancerous tumor and 84% for cancerous tumors with an AUC of 0.798. Feature extraction could be automated in the future by implementing deep learning techniques.

Table 2.3: Summarization of distinguished work on breast carcinoma diagnosis using Machine Learning Classifiers

Data Set Used	**Method Adopted**	**Performance**	**Analysis**	**Reference**
2301 Patients (Private Dataset)	Support vector regression(SVR) with RBF, Linear and Poly kernel, Decision Tree Regression (DTR) model, Stochastic Gradient Descent model	The best accuracy is given by SVR with Linear kernel and DTR model	• SVR linear model and DTR model best suited for survival time prediction. • Tumor-integrated clinical feature (TICF) performs better than the Nottingham prognostic index (NPI).	[7]

SEER [45]	NB, LR, DT	Survivability prediction on all stages together can cause overestimation of performance	• Distant summary stage shows the worst performance due to lesser instances of training data • Compared to the joint model, the stage-specific model yields superior results.	[8]
WBCD [40]	Hybrid of K-means with SVM (K-SVM)	97.38%	• Reduces training time by reducing input features from thirty-two to six • Feature selection during training and validation phase not required	[58]
	SVM, Decision Tree-C4.5, Naive Bayes, KNN	97.13 %	• SVM is manifested to be the finest classifier with the least error rate and highest correctly classified instances on WBCD	[59]
UCI machine learning repository	SVM, DT-J48, NB, LR	**Dermatology-** NB+LR (97%) **Breast Cancer-** SVM-PCA (97%) **Chronic Kidney Disease-** J48 (99%)	• High accuracy • Accuracy could be improved by considering more data and parameters such as age, sex, locality • CNN could be used in the future to automate	[60]

			feature extraction.	

2.5 Methodological Work On Breast Carcinoma Diagnosis Using Deep Learning Strategies

This section outlines the contributions of researchers in recent years for diagnosing breast carcinoma through deep-learning techniques[13].

Al-antari et al. [61] drafted a CAD model on the In-Breast database and implemented the "You Look Only Once (YOLO)" model for detection. The authors proposed a "full resolution convolutional network (Fr-CN)" and a DCNN for segmentation & categorization, respectively. Compared to other CAD models, the presented method attains significant performance with a testing time per image of only 12.23s. The network attained an accuracy of 95.64% after the detection, segmentation, and classification of masses with an AUC of 0.94. In [62], the experts utilized YOLO on the DDSM dataset. The outcomes demonstrate that YOLO performs exceptionally well when it comes to mass identification over dense tissues and pectoral muscle. A one-stage object detector, called Retina Net, is the foundation of another mass detection model put forth by experts in [63] is based on one-stage object detector, i.e., Retina Net. The algorithm provided a greater sensitivity value concerning false positives per image (FPPI). The model was implemented on a publicly available dataset, i.e., In breast, and an in-house dataset, i.e., GURO. The authors concluded that performance could be improved by integrating large datasets using transfer learning. Ting et al. [64] utilized an "Interactive detection-based lesion locator (IDBLL)" and proposed an improved CNN model for diagnosing breast malignancy. The authors attained an accuracy value of 90.50% on the MIAS dataset.

In [65], the authors presented an efficient approach for classifying mass lesions on BCDR-FM using a CNN. The stated framework attains an AUC value of 0.826 and outperforms other cutting-edge strategies by combining learned and hand-crafted features. Their work could be continued in the future with high-resolution images provided by BCDR-DM. In [66], the researchers implemented Convolutional Neural Network Discrete Wavelet (CNN-DW) and Convolutional Neural Network Curvelet

Transform (CNN-CT) on the IRMA dataset. Results have shown that using SVM instead of the Softmax layer can improve classifier performance. CNN-CT outperforms CNN-DW with an accuracy value of 83.74, 83.11, and 81.49 using the SVM layer, SVM layer with 10-fold cross-validation, and Softmax layer. Their performance can be enhanced by applying different CNN architectures and various deep learning techniques in the future. H. Li et al. [67] proposed DenseNet II architecture for classifying mammograms. The authors substituted the first convolutional layer with the inception framework. The proposed model shows significant performance when analyzed with other pre-trained networks and attains an accuracy of 94.55% and an AUC of 0.91. Therefore, this design improves speed and efficiency and solves pre-trained models' gradient descent by reducing parameters and enhancing feature reuse. In [68], the authors proposed a supervised approach for evaluating percentage density (PD) and compared the DCNN approach with the statistical learning approach. PD is referred to as the proportion of dense area to the breast region on digital mammograms. They concluded that, for PD evaluation, the DCNN method is superior and powerful to a feature-based learning approach. In the future, their system can be extended by including mammograms from different vendors, using various imaging modalities, and implementing transfer learning.

Pectoral Muscles hinder the detection process and should be removed before diagnosis. Mughal et al. [69] proposed a hybrid technique to eliminate pectoral muscle using a curve stitching technique. The adaptive hysteresis thresholding segmentation technique identifies the ROI and segments the breast region of a digital mammogram. Results have shown that this method obtains the loftiest recall of 96.6% on MIAS and 96.4% on the DDSM dataset when contrasted with other revolutionary approaches. FFDM offers low detection sensitivity, and MRI, on the other hand, provides high costs with longer scanning time. "Contrast-Enhanced Digital Mammography (CEDM)" combines the advantages of both, offers four times faster scanning, and is only 1/6 the price. CEDM generates low-energy images similar to FFDM and recombined images comparable to breast MRI [70]. For improving the diagnosis of breast carcinoma, "SD-CNN (Shallow-deep CNN)" was proposed by authors in [70] by utilizing CEDM. The authors used deep CNN to compare the results obtained through low energy(LE) images and by the blend of LE and recombined images.

Secondly, due to the restricted availability of CEDM, the authors applied shallow CNN (S-CNN) and obtained a 'virtual' recombined image from FFDM images. After evaluation, with virtual recombined imaging features, results show an increase in the model's performance.

Breast calcifications are minute deposits of calcium on the breasts. They commonly appear as white specks or dots on the mammograms. Calcifications are generally non-cancerous, but micro-calcification clusters could indicate a warning sign of malignancy. The authors in [71], utilized hand-crafted and deep learning features from digital mammograms to perform breast micro-calcification diagnosis using DCNN. Using filtered deep features, 86.89% recall and 89.32% precision was obtained. Researchers concluded that although deep features surpass hand-crafted (morphological) features, on the other hand, hand-crafted features could guide CNN to attain higher accuracy by providing additional information.

On the other hand, authors in [27] classify multi-view mammograms and segmentation maps using pre-trained deep learning models. As opposed to classifying individual lesions, this holistic approach classifies the whole mammogram comprising "Craniocaudal (CC)" and "Mediolateral oblique (MLO)" views. The findings reveal that this method produces good results on the In-Breast and DDSM datasets, with an AUC of 0.9. Arefan et al.[72] examined the results using CC and MLO views on mammography patients and estimated that the CC perspective shows superior performance than the MLO view. The authors also observed an increase in AUC value when the two views are used in combination for heavy breasts.

In medical imaging, Multiple instance learning (MIL) provides room for weakly labelled data and reduces the effort and cost of marking all the dataset images. Instead of taking a group of labelled instances, it considers the set of labelled bags comprising multiple instances. P.J. Sudharshan [73] used MIL and implemented several methods on the Break His database. The authors proposed two ways to implement the process. The first method considers patients as bags and instances are the patches extracted from images, and the second assumes images as bags and instances are the patches. Non-parametric MIL and Multiple Instance Learning Convolutional Neural Network(MILCNN) attain significant results with 92.1% accuracy. The authors

concluded that MIL outperforms other methods for classifying cancer on histopathology images.

Similarly, Samala et al. [74] concluded that when training samples from one modality are inadequate, multi-task learning is a competent technique for implementing deep networks in medical image processing. Another work on MIL was proposed in [75] for mass detection in DBT images. The researchers developed a framework based on deep CNN(DCNN)) with multiple instances learning random forest approach(MILRF), i.e., DCNN-MILRF and compared the presented structure with two other proposals- first based on hand-crafted features, i.e., MIL-CAD and second based on Deep Cardinality Restricted Boltzmann Machine (Deep CaRBM). Findings indicate that the DCNN-MILRF performs superior than the other two approaches with 86.8% accuracy. In the future, their work could be explored with a larger DBT dataset by combining 3-D imaging information with 2-D information.

In [76], the authors proposed a model for lesion detection on ultrasound images on two datasets, A&B with 306 and 163 images. The authors used three DL approaches, patch-based Le-Net & U-Net, and transfer learning-based Fully Convolutional Network-Alex Net (FCN-Alex Net). Results show that FCN-Alex Net gave the best performance on Dataset A and patch-based Le-Net on Dataset B in terms of F-measure and false-positive/image but at the expense of longer training sessions. Their work could be extended by increasing training data and implementing segmentation and classification on ultrasound lesions. Another similar work on ultra-sonographic images was proposed in [77]. The experts utilized a deep CNN and developed two architectures called "Mt-Net" and "Sn-Net" to categorize tumors in the breast lesions. Results reveal that the presented approach surpasses other revolutionary methods and can be elaborated in the future by combining several imaging modalities. Table 2.4 summarizes research on breast carcinoma diagnosis using deep learning models.

For training CAD models using deep learning techniques, training information is needed in bulk. However, the collection of the labelled dataset is a challenging job in the health sector. Different augmentation strategies can be applied to increase dataset size. Due to the extensive availability of a huge amount of images, numerous researchers have utilized the Break-His dataset in recent years to classify histopathology images. By analyzing the outcomes independently for patch-wise and

image-wise classification, the authors of [78], [79] have achieved remarkable success in employing histopathological pictures to classify breast carcinoma. Using multiple pre-trained networks, including Res-Net50, Alex-Net, and VGG-16, [80], [81] explored histological image classification and achieved excellent outcomes. In another work, Brancati et al. [82] utilized residual Convolutional Auto-Encoder (CAE) i.e. Fusion-Net, to multi-classify lymphoma in histology pictures. Supervised Encoder Fusion-Net (SEF) is an acronym of the supervised classification method that the authors employed in addition to the first method, classification by reconstruction, which involves training CAE in an unsupervised way. The authors compared results obtained using SEF with other existing deep learning networks (Fusion-Net, U-Net, Res-Net) under the same conditions and on the same datasets. They deduced that SEF offers excellent outcomes for the situations in which it was employed and can be tested for other histopathological image assessments in the future.

A novel framework was proposed in [83]. The Squeeze and Excitation Residual network (SE-Res Net) module was designed to lower training parameters and guard against overfitting of the model. The stated methodology attained an accuracy of 90.66 to 93.81%. Their work can be elaborated to study cell duplication and irregular distribution of color in pathological images in the future. Three models were presented in [84] for the histopathological image classification: CNN, Long-Short-Term-Memory (LSTM), and a CNN/LSTM hybrid. Using K-Means and Mean shift, the scientists were able to extract statistical and unobserved structural information from the data. With a 91% accuracy rate, CNN with Softmax classifier performs better than the other two approaches. Through the combination of pathological pictures and structured information from the Electronic Medical Record (EMR), the researchers in [85] developed a hybrid DL model. The fusion of both high and low-dimensional information resulted in a substantial improvement in diagnostic accuracy. Another paradigm on the Break His dataset was proposed by Sharma and Mehra [86]. Firstly, the authors used conventional classifiers to carry out classification using manually created features. Secondly, pre-trained networks were employed as feature extractors. Findings demonstrate that the pre-trained model-based strategy performs better than the hand-crafted and baseline approaches. The finest accuracy was accomplished by the VGG-16. Their work could be elaborated further to perform

layer-wise fine-tuning of the model and implement an ensemble model with a pre-trained network.

Bevilacqua et al. [87] perform a supervised classification of breast tomosynthesis images by comparing shallow and deeper neural networks. The authors proposed two frameworks: the first framework was based on hand-crafted (morphological/textual) features, i.e., classification by ANN, and the second was a non-neural classifier (i.e., automatic feature extraction by CNN) in which different CNN models (GoogleNet, ResNet, AlexNet, and VGGNet) were evaluated. Results evinced that the CNN-based model outperforms the ANN approach. In [88], the authors classify mammograms and DBT images on an in-house dataset by evaluating ten different models using pre-trained Alex Net. In their research, 2-D mammography outperforms 3-D tomosynthesis models in classifying cancer. However, 3-D tomosynthesis has proven to be robust in manual cancer detection. The variation in result might be due to memory limitations; as a result of which the authors considered the only subset of 3-D frames, and thus, significant information loss might have occurred in 3-D pictures. For achieving optimal performance, their research could be extended in the future for developing an ensemble classifier integrating 2-D mammography with 3-D tomosynthesis images. [89] stated a framework for mass detection in DBT images and showed a comparison between DCNN and feature-based CAD. Unlike other research, where pictures of natural scenes were utilized for pre-training a DCNN, the authors first used mammography images for training a DCNN, and then this pre-trained model was used for mass detection in DBT pictures using transfer learning. Results have shown that the AUC value increases from 0.81 to 0.90 after transfer learning with DBT. Secondly, DCNN-based CAD surpasses feature-based CAD and thus improves sensitivity from 83% to 91%. By applying pruning, the number of neurons and parameters were lowered by 87.2% and 34.4%, respectively. Another work on mass detection on DBT images was proposed in [90] using a "Faster Region-Based Convolutional Neural Network(Faster-RCNN)." The authors modified the DCNN proposed in [76] and compared it with the RCNN. Results show that RCNN outperforms DCNN with an AUC value of 0.96. The false-positive rate is diminished, but for the sake of cost and speed, i.e., RCNN is sluggish and costly as compared to DCNN.

In [91], the experts utilized gene expression datasets to anticipate clinical outcomes of breast carcinoma. The classification task consists of 2 phases- Unsupervised feature learning, i.e., combining PCA and auto-encoder, and, second, supervised classifier learning, i.e., constructing an ensemble classifier based on the Ada-Boost algorithm. Based on experimental findings, the suggested strategy, i.e., PCA-AE-Ada-Boost, gave an AUC of 0.714 and an accuracy of 0.85. Despite its advantages, the authors highlighted some limitations- firstly, difficulty identifying essential features for prediction tasks; secondly, increasing the generalization capability of the proposed method by including more publicly available datasets. Deep learning methods can also be utilized for extracting extensive clinical information for breast carcinoma from multiple categories of clinical notes in Chinese [92]. In [92], the authors proposed a system with two constituents- Named Entity Recognition (NER) and Relation recognition component. The bi-directional Encoder Representations from Transformers (BERT) was fine-tuned for extracting the notion and features from clinical breast cancer documents. Results have shown that fine-tuned BERT outperforms traditional Bi-directional long-short-memory-conditional random fields (Bi-LSTM-CRF) algorithms. The approach yielded precision, recall, and f-measure of 0.927, 0.939, and 0.935 for NER and 0.976, 0.959, and 0.967 for relation extraction, respectively. For information extraction, the proposed approach considered only 100 patients; therefore, additional data annotation could be done to increase dataset size in the future. Secondly, the proposed model could be enhanced to integrate data from alternative imaging techniques like computed tomography (CT), MRI, and pathological microscopic scans.

In [67], the authors presented an improved and effective neural network model called DenseNet II for diagnosing and classifying breast carcinoma. The model yields 94.55% accuracy and 0.91 AUC. Another research[87] compared two approaches, i.e. ANN and CNN for the categorization of breast tomosynthesis images. The research concluded that the CNN-based approach gives higher accuracy and AUC value than other approaches. The authors in [93] utilized the Break His Dataset and performed breast cancer classification on histopathology images using VGG-16, VGG-19, Mobile Net, and ResNet-50. VGG-16 outperforms the other models with the greatest accuracy of 94.67%. [94] presents a systematic literature survey on ML and DL

approaches for diagnosing breast carcinoma using mammograms. The study discusses the imaging modalities, datasets, and, techniques used, for the breast cancer CAD system. Performance measures, potential limitations, and future challenges are also outlined. In [95], the researchers have proposed a two-stage transfer learning approach for diagnosis on ultrasound pictures. In the first stage, cancer cell line images were utilized, and the features learned from this stage were transferred to categorize benign and malignant carcinoma in the ultrasound pictures. The authors attained 97.8% accuracy on the MT-Small (Private) Dataset. [96] provides an in-depth study on analyzing medical images using ML and DL techniques. The authors observed that DL outperforms ML models when it comes to evaluating enormous quantities of data. This study focuses on the detection & diagnosis of various medical illnesses such as Breast tumors, Brain disease, Diabetes, etc.

The researchers in [97] presented a multi-activation deep neural network for diagnosis on the WBCD dataset. The authors concluded that the multi-activation proposed deep neural network model performs better than other single-activation networks. In the future, researchers can use different combinations of activation functions for deep neural networks. Senan et al. [98] used Alex Net on the Breast cancer digital repository (BCDR) dataset to categorize malignancy on histopathology images. The proposed approach gave superior results to previous models, with an accuracy of 95% at the magnification factor of 40x and 400x.

Table 2.4: Summarization of distinguished work on breast malignancy using Deep Learning Strategies

Imaging Modality/ Dataset Used	Method Adopted	Performance	Analysis	Reference
Mammograms (Labeled and Unlabeled)	CNN with Semi-Supervised Learning (SSL) algorithm	Accuracy: 82.4 AUC: 0.88	• Best result obtained using mixed labelled and unlabelled data • Useful when the labelled data is not available in bulk • Performance of the system increases using unlabeled data	[99]
Histopatho-logical images	Inception network, DCNN, Gradient Boosting Tree classifier	Accuracy: 96.4	• Gradient Boosting tree algorithm performs well even in imbalanced training data and limited samples, thus avoiding over-fitting. • Improves performance when DCNN is fused with gradient boosting tree	[100]
	Pretrained Xception Network	Pre-eminent performance	• Multi-task CNN outperforms single task CNN. • Fine-tuned CNN outperforms CNN trained from scratch • Inclusion of prior knowledge in feature extraction step overcomes	[101]

			intra-class variance problem in pathological images.	
SFM, DM, and DBT	Two-stage transfer learning	ROI-based & View-Based AUC Single-stage: 0.84 and 0.85 Two-stage: 0.90 and 0.91	• A supplementary phase of transfer learning from a similar related domain provides an improvement in learning. • Hyper-parameters were not optimized due to smaller datasets and less computational cost. • Did not compare performance with conventional feature engineering methods	[102]
FFDM, US, Dynamic Contrast-Enhanced MRI(DCE-MRI)	VGG-19 model	AUC (Max pooled features): FFDM: 0.81 US: 0.87 MRI: 0.87 AUC (Fusion classifier): FFDM: 0.86, US: 0.90, MRI: 0.89	• Max-pooled features performed better as compared to fully connected features with or without pre-processing • Fusion classifier outperforms CNN and convolutional CAD classifier	[103]

2.6 Related work on the Generalization of Deep Convolutional Neural Networks

Generalization error is another critical parameter that needs to be considered when using transfer learning with limited training data. DCNN can memorize the training data due to millions of parameters and trainable weights[104]. Thus, memorization, over-fitting, and a classifier model with enormous specifications may result in generalization errors [105].

Samala et al. [105] considered the impact of learning capacity and transfer learning on the generalization error by plotting ROC curves as an evaluation parameter. The noise was introduced by corrupting image labels and the input image. The authors evaluated the correspondence between learning and memorization of deep networks by freezing convolutional layers of Alex-Net and Google-Net. Findings show that generalization error increases when training is done with noisy data. The authors also concluded that a proper strategy should be chosen based on available training data to minimize overfitting and improve the network's generalizability when implementing transfer learning. [106] presents Domain Generalization for deep learning. Using several Deep Learning architectures, a Single Source Domain Generalization(SSDG) approach is presented for breast carcinoma diagnosis in mammograms. The authors concluded that models based on transformers were more resistant to domain changes and SSDG techniques minimize the domain shifts and thus enhance the model's efficiency in unknown clinical contexts. Authors in [107], show the comparative view of several types of noises in digital image processing. They have proposed noise models and summarized that noise affects produce artifacts such as faulty edges, invisible lines, corners, and fuzzy objects. Various filtering techniques could be applied for noise removal. Deep Neural models should be trained on these images to find the model's generalizability.

[108] provides a review of Generalization errors in Deep Learning. The authors summarized that Regularization techniques lower the model's complexity and thus reduce generalization error. The authors in [109] proposed a framework for Impulse Noise Detection using a Modified Robust Outlyingness Ratio (mROR) on mammogram images. This study could be modified further for the deployment of real-time clinical images. In [110], the authors determine the vulnerability of deep learning models on

medical images through three attacks i.e. Fast Gradient Sign Method(FGSM), Projected Gradient Descent(PGD), and Basic Iterative method(BIM) with different perturbation levels. Medical DL models are more vulnerable to adversarial assaults resulting in decreased accuracy values. Thus, it was concluded that Adversarial Training increases the model's robustness for different attacks. Another approach to domain generalization is proposed in [111], based on a multi-view contrastive learning technique. The outcomes show the model's efficiency in both seen and unseen domains. Thus the method could be used for diagnosing lesions in mammograms.

2.7 Chapter Summary

In this chapter, an overview of recently developed, peer-reviewed CAD systems that employ ML and DL techniques for the diagnosis of breast malignancy is presented. These systems are compared with formerly authorized approaches. Technical specifications along with the benefits and drawbacks of each framework are explained. Different strategies for implementing Deep Convolutional Neural Networks are also discussed. AI has expanded over the last ten years, and AI-based applications in the medical field have demonstrated positive results at lower expenses and with greater effectiveness.

Previous research findings elucidated that when the dataset is broad, DL performs better for diagnosing breast cancer than traditional ML. To improve healthcare over the long term, practical and scientific investigation is desperately needed. Consequently, to improve survivability and lower mortality rates over time, early diagnosis and prognosis have become essential. By aiding radiologists in their examination of medical images, new artificial intelligence tools are improving the prognosis of cancer patients.

.

PUBLICATION

The research discussed in this Chapter has been published in the following journal:

- G. Chugh, S. Kumar, and N. Singh, “Survey on Machine Learning and Deep Learning Applications in Breast Cancer Diagnosis,” Cognitive Computation. 2021, doi: 10.1007/s12559-020-09813-6. **(SCIE Indexed, IF=4.3)**

CHAPTER 3

TRANSNET FRAMEWORK: A DUAL APPROACH FOR BREAST CARCINOMA DIAGNOSIS

3.1 Overview

Breast Carcinoma is a deadly disease; therefore, timely diagnosis is one of the most critical concerns that must be addressed globally since it can significantly enhance overall survival rates. Currently, Medical Imaging relies on ML and DL for accurate and early recognition of sickness. In this chapter, an architecture is stated for diagnosing & classifying breast tumors using deep learning approaches. Two experiments were performed on the CBIS-DDSM dataset. In the first approach, i.e., Deep feature extraction with ML classifier head, DCNN models such as VGG-16, VGG-19, ResNet-50, and ResNet-152 are deployed as feature extractors, and the obtained features are utilized for training conventional machine learning classifiers. The second approach, called Deep Learning feature extraction with a neural network classifier, exploits Mobile Net, VGG-16, VGG-19, ResNet-50, ResNet-152, and, DenseNet-169 for feature extraction and categorization.

3.2 Data Set Utilized

In Chapter 1, we discussed several imaging techniques for diagnosing breast carcinoma such as Mammography, MRI, Ultrasound, DBT, Histopathological pictures, etc. The details, along with the pros and cons for each modality, could be referred to in [13]. Breast cancer diagnosis via mammography is still recognized as a benchmark, as it detects tumors at initial stages and is associated with low radiation compared to other modalities. In this research, we have used a mammography dataset i.e., CBIS-DDSM [112], which contains 9648 mammography images for 2412

patients. 695 cases are Normal, 867 are malignant(Positive), and 850 are benign(Negative). After applying data augmentation (Discussed in Section 3.4.2), the dataset comprises 55,885 images which include: 16,103 Normal, 20,088 Malignant(Positive), and 19,694 Benign(Negative) samples. We have utilized the data maintained as a tfrecords file for TensorFlow.

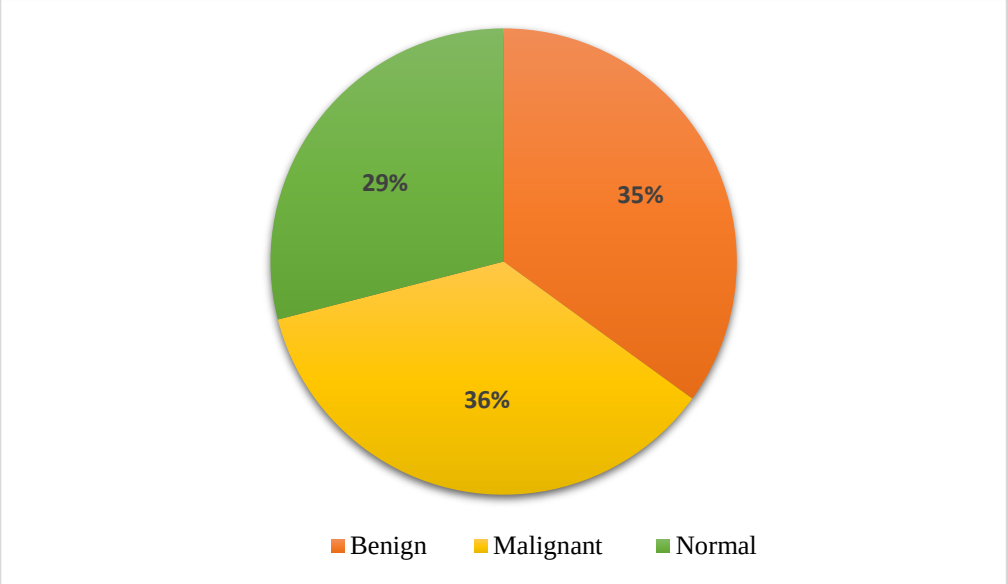

Fig. 3.1: CBIS-DDSM Data Set [112]

3.3 Deep Convolutional Neural Networks(DCNN) & Machine Learning(ML) Classifiers

The proposed Transnet Framework is a dual approach for diagnosing breast cancer. The framework utilizes various DCNNs and ML Classifiers.

3.3.1 Deep Convolutional Neural Networks(DCNNs)

The specific details of DCNN architectures implemented in the proposed framework are discussed as follows[113]:

(i) **VGG Net** [114]: VGG Net is a very popular and efficient DCNN that has shown promising results in image classification. In VGG Net, 11 x 11 and 5 x 5 filters

are substituted with a stack of 3 x 3 filters. The simultaneous deployment of a 3 x 3 filter might induce the impact of a large filter size (7 x 7, 5 x 5). VGG Net places 1x 1 convolution in between the convolution layers. Owing to the abundance of parameters, VGG Net is computationally expensive and requires longer training time on the system with less computational power. In this framework, we have utilized VGG-16 and VGG-19 with 16 layers and 19 layers of depth. The original layered VGG-16 and VGG-19 structure is depicted in Figure 3.2.

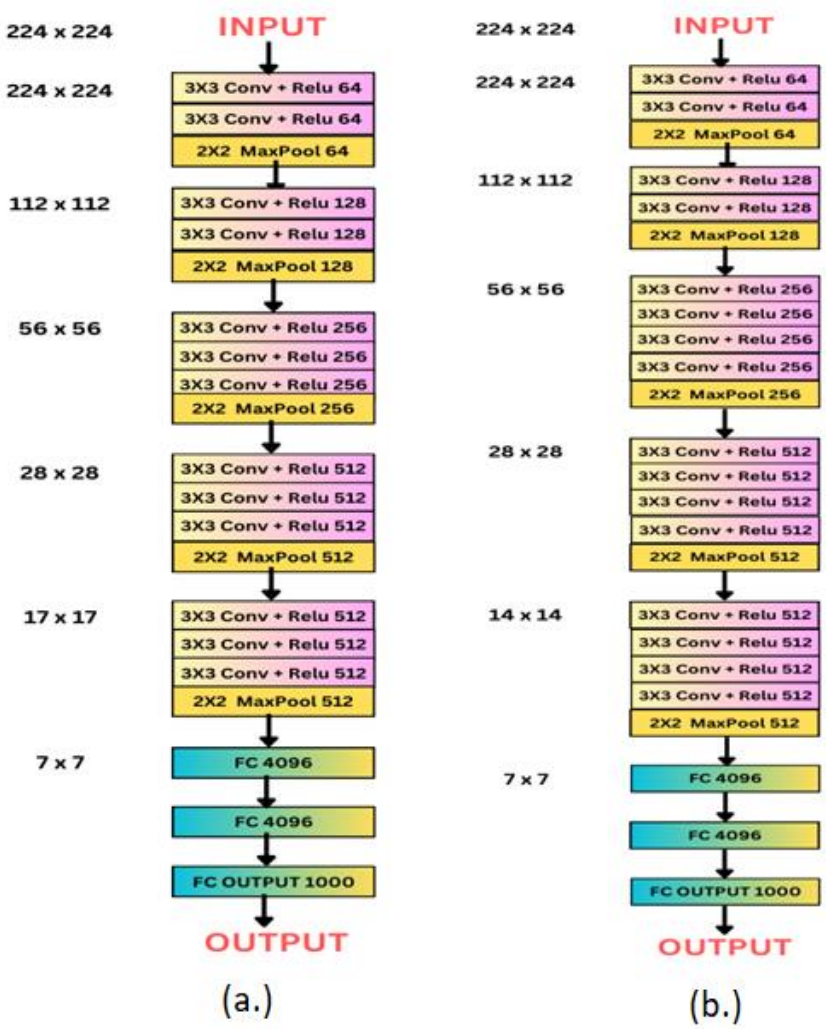

Fig. 3.2: VGG Structure[114], [115] (a.) VGG-16 (b.) VGG-19

(ii) Res Net [116]: This is a renowned deep-learning pre-trained network, also called as Residual Network. As layers increase in a network, a problem known as a

vanishing gradient usually occurs. To overcome this, Res Net introduced the concept of residual learning. Here, we use Skip Connections, i.e., it bypasses a few stages of training and connects immediately to the output. Res Net comes in different versions with 50/101/152 layers' depth with an input image size of 224 x 224. We have utilized two versions of ResNet i.e. ResNet-50 which has 24,649,953 trainable parameters and ResNet-152 which has 59,303,265 trainable parameters. The original layered architecture of both is shown in Figures 3.3 and 3.4.

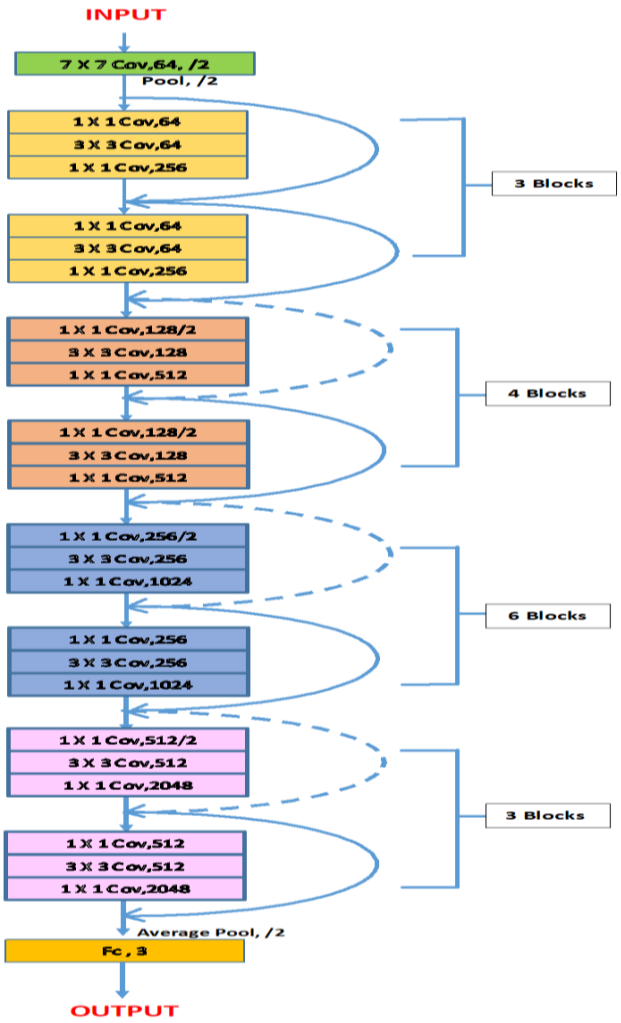

Fig. 3.3: ResNet-50 Architecture[117]

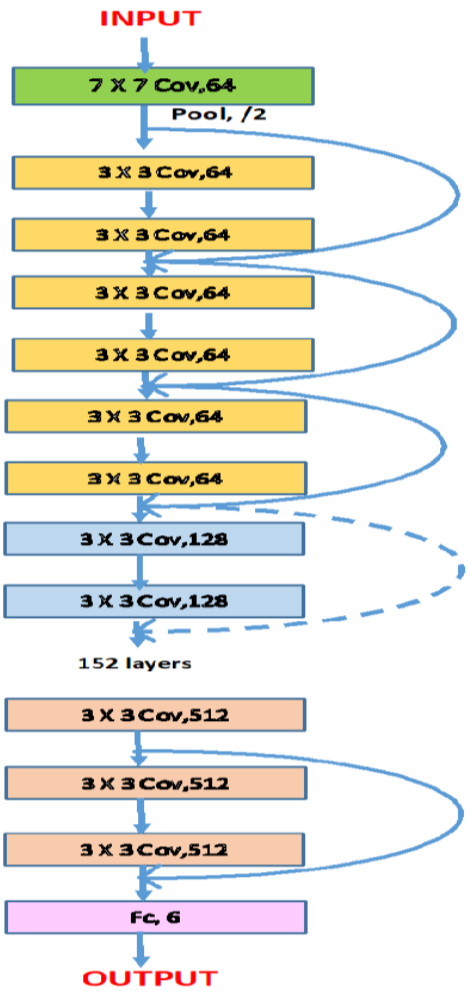

Fig. 3.4: ResNet-152 Architecture[118]

(iii) Mobile Net [119]: This model is developed specifically for mobile applications. Mobile Net reduces the computational complexity, i.e., the number of parameters, by using depth-wise separable convolutions. It has 3,521,569 trainable parameters. These low-powered small models are suitable for applications where resources are limited. Figure 3.5 depicts the Mobile Net architecture.

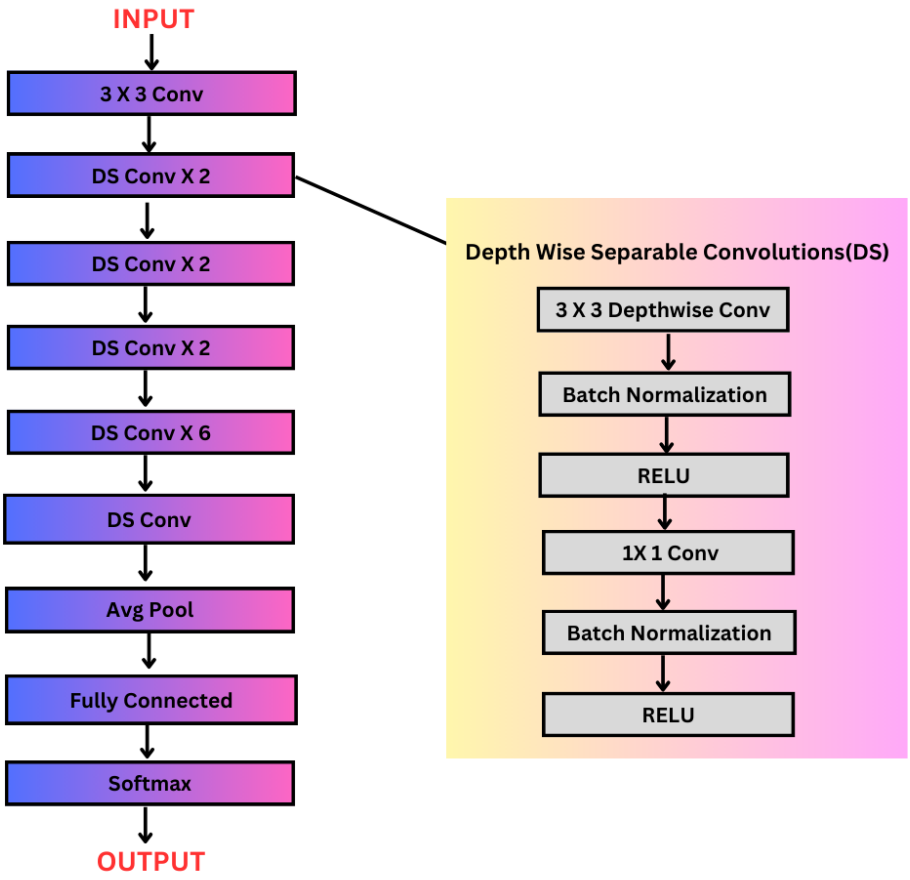

Fig. 3.5: Mobile Net architecture[119], [120]

(iv) DenseNet [121]**:** Also called Dense Convolutional Network. Dense Nets provide numerous benefits, such as solving vanishing-gradient problems, enhancing feature propagation, reusing features, and reducing parameter count. Dense Net models are easy to train because they provide enhanced information flow across the network. Dense Nets usually have hundreds of layers and provide no optimization problems. These networks have several versions- DenseNet-121, DenseNet-169, DenseNet-201, etc. In this research, we have utilized the DenseNet-169 model which possesses a large number of trainable parameters i.e. 12,994,913, and thus could be deployed as feature extractors in various computer vision tasks. The last layers of the network are modified as per the requirement. Figure 3.6 depicts 169 layers' architecture of DenseNet-169 with 4 dense blocks and 3 Transition Layers.

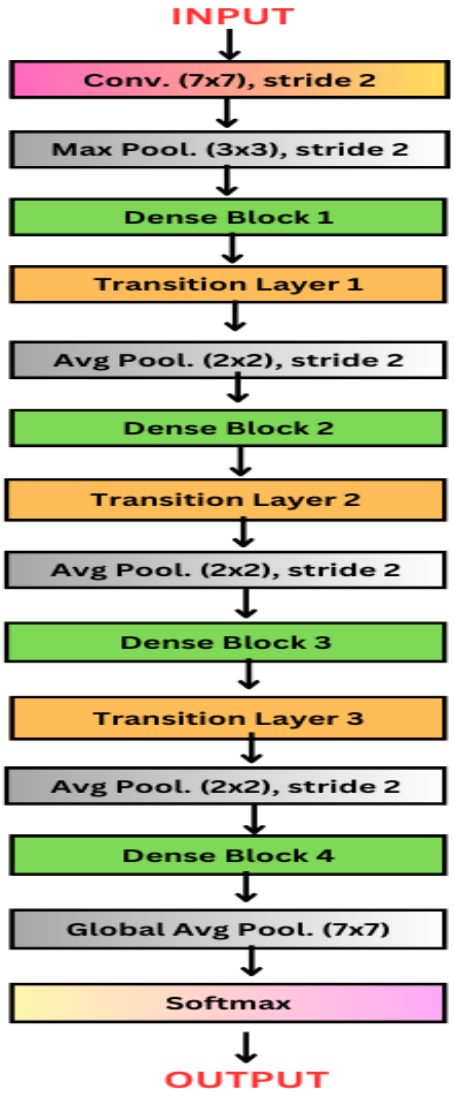

Fig. 3.6: DenseNet-169 architecture[122]

3.3.2 Machine Learning Classifiers

Following ML classifiers have been deployed in the proposed framework:

(a.) **k-Nearest Neighbour(kNN)**[123]**:** This is a classification method that is widely used for disease prediction. One variable parameter in the algorithm is called k, and it represents the number of "nearest neighbours". Data points closest to the query point are determined based on their closest distances. To determine which class appears the most, it locates the k nearest data points

and then uses a majority voting process. Because of its simplicity, the KNN algorithm is one of the most used machine learning algorithms for classification problems.

(b.) **Support Vector Machine (SVM)**[124]**:** SVM establishes a decision boundary between two classes so that one or more feature vectors can be used to forecast labels. The hyperplane, which represents this decision boundary, is oriented to be as far away as feasible from the nearest data points for each class. Support vectors are the locations that are closest together. Regarding identifying minute patterns in complicated datasets, SVM is far more effective than other ML techniques. A function called the SVM kernel converts low-dimensional input space into higher-dimensional space. Thus, the kernel function could speed up some computations that would otherwise need high-dimensional space calculations. Different kernel functions could be utilized- Linear, Polynomial, Gaussian, and Sigmoid Kernels.

(c.) **Random Forest (RF)**[125]**:** This method employs ensemble learning i.e. it merges the outcomes of multiple classifiers and determines the average. Ensemble classification yields better results using many classifiers than individual classifiers alone. A voting procedure is then implemented to select the class tag for instances without labels. To ascertain the category name of an unlabelled instance, each ensemble tree functions as a base classifier. Majority voting is employed in which each classifier casts one vote. The instance is then categorized by selecting the class tag with the maximum vote count.

(d.) **Adaptive Boosting (Ada Boost)**[126], [127]**:** This is an ensemble learning method, useful for a range of classification and regression applications. By merging various weak learners (such as decision trees) into a powerful learner, this supervised learning system can classify data. In Ada Boost, the training dataset's occurrences are weighted according to how accurately they were previously classified. This algorithm is adaptive because it adjusts

weaker learners thereafter to prioritize cases that earlier classifiers misclassified. Being sensitive to outliers, Ada Boost is not appropriate for noisy data.

(e.) Extreme Gradient Boosting(XGB)[128], [129]**:** The approach of enhancing a particular weak model by integrating it with several other inadequate models to produce a final strong model is known as "gradient boosting". XG Boost is a flexible and extremely precise gradient-boosting model. XG Boost builds trees in parallel. It employs a level-by-level strategy for examining gradient values and then assessing the split quality at each potential split point in the training data. This algorithm applies regularization methods to lessen overfitting and enhance the generalization of the model.

3.4 Proposed Transnet Framework

We have presented a CAD framework, i.e., Transnet for diagnosing and classifying breast carcinoma using Deep Neural Networks. Transnet presents a dual approach for breast carcinoma diagnosis. It involves deep feature extraction using Deep Convolution Neural Network (DCNN) Models and then their classification using Machine learning classifiers and Neural Net classifiers respectively. Pre-processing and augmentation of the CBIS–DDSM dataset were performed to optimize the proposed strategy and to reduce overfitting. Several measures, including Accuracy, AUC, Precision, Recall, and Loss, are utilized and plotted on graphs to estimate the models' efficiency.

3.4.1 Pre-Processing

An important factor affecting Deep Convolutional Neural Networks' performance is Pre-Processing. The steps followed for pre-processing and augmentation are discussed as follows[113] and shown in Figure 3.7.

(i) Min-Max Normalize

Normalization is scaling your data within a specified range. Prediction of Deep Neural Networks could be further enhanced through intensity normalization as it reduces data variability[130]. Thus, the intensity value of the pixels in the pictures is scaled down in the range (0,1). This process also reduces the unwanted noise in the image.

(ii) CLAHE

To boost the contrast of greyscale pictures, Contrast Limited Adaptive Histogram Equalization(CLAHE) is applied. It works in two steps: firstly, enhancing the contrast of pictures and secondly, limiting the contrast within a range to prevent the image's noise from being amplified. This strengthens the network's potential to acquire minute details, textures, and characteristics from the mammogram [131]

(iii) Padding

The process of adding more data to the image borders is known as padding. Thus, to change a convolutional layer's output size, padding is a commonly used technique in CNNs. Without padding, the output size will be smaller since convolutional filters won't process the input borders. Padding allows for the preservation of the input size; before the convolution, we add a border to allow for the processing of the original border[132]. The majority of computer vision tasks accept square images as input. In this step, images are padded into squares to feed them into the model.

(iv) Centre-Crop

This operation crops the core area of the image. This will allow the DCNN model to identify the irregularity with fine details. Every image in the dataset was scaled down to 299 X 299.

3.4.2 Data Augmentation

DNNs need substantial amounts of training information. Medical Images are not accessible in abundance, and thus problems such as overfitting might occur. Augmentation involves enlarging the dataset, i.e., the creation of supplementary data from the existing through several operations such as translation, rotation, scaling, etc.

[119]. In the proposed approach, we applied the following two operations on the dataset:

(a.) Random Flipping

The images are randomly flipped both horizontally and vertically.

(b.) Random Rotation

In this case, images are randomly rotated at different angles.

After applying data augmentation, the dataset comprises 55,885 images which are then split into three ratios i.e., 80:10:10 for the training, validation, and test sets.

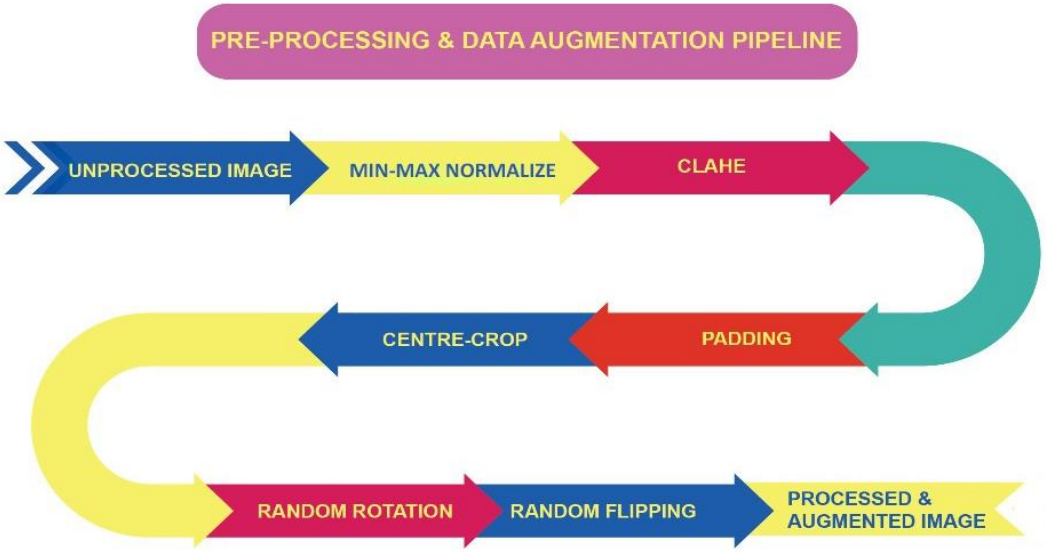

Fig. 3.7: Pre-Processing Pipeline

3.5 Model 1: Deep Feature Extraction with Machine Learning Classifier

3.5.1 Proposed Architecture

This strategy employs the following DCNNs i.e. VGG-16, VGG-19, ResNet-50, and ResNet-152 as feature extractors. Extracted features are utilized for training the Machine Learning classifiers. This strategy deploys the following classifiers: KNN

with a K value of 8, SVM with RBF (Radial Basis Function) kernel, RF, Ada Boost, and XGB model. Figure 3.8 shows the proposed architecture for Model 1[113].

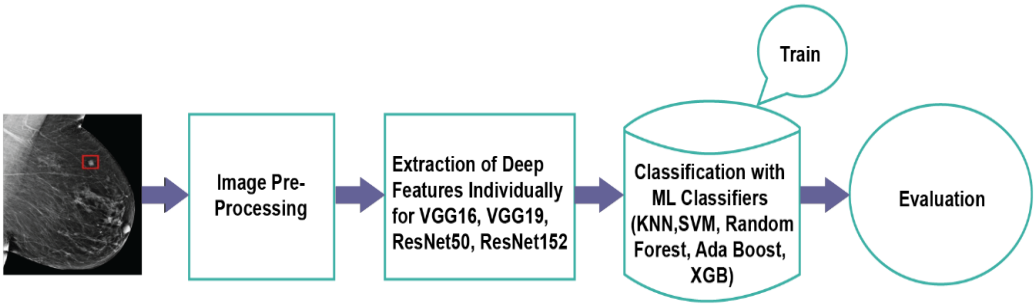

Fig. 3.8: Deep Feature Extraction with ML classifier

3.5.2 Evaluation Metrics

In this approach, Deep feature Extraction with ML classifier three performance measures have been evaluated: Accuracy, Precision, and Recall. Tables 3.1 to 3.3 show the corresponding evaluation metrics[113]. These results are analyzed in section 3.7.

Table 3.1: Performance metrics on the Training dataset (Model 1)

S No.	DCNN Model	Classifier	Accuracy	Precision	Recall (Sensitivity)
1	VGG-16	KNN	0.90	0.90	0.90
		SVM	0.94	0.94	0.94
		RF	1.00	1.00	1.00
		Ada Boost	0.90	0.88	0.89
		XGB	1.00	1.00	1.00
2	VGG-19	KNN	0.89	0.9	0.89
		SVM	0.93	0.93	0.93
		RF	1.00	1.00	1.00
		Ada Boost	0.90	0.89	0.9
		XGB	1.00	1.00	1.00
3	ResNet-50	KNN	0.91	0.9	0.91
		SVM	0.95	0.95	0.95
		RF	1.00	1.00	1.00

		Ada Boost	0.91	0.9	0.91
		XGB	1.00	1.00	1.00
4	**ResNet-152**	KNN	0.91	0.9	0.91
		SVM	0.98	0.97	0.97
		RF	1.00	1.00	1.00
		Ada Boost	0.92	0.94	0.93
		XGB	0.98	0.98	0.98

Table 3.2: Performance measures on the Validation dataset (Model 1)

S No.	DCNN Model	Classifier	Accuracy	Precision	Recall (Sensitivity)
1	**VGG-16**	KNN	0.88	0.86	0.88
		SVM	0.91	0.91	0.91
		RF	0.89	0.88	0.89
		Ada Boost	0.89	0.87	0.89
		XGB	0.91	0.90	0.91
2	**VGG-19**	KNN	0.88	0.87	0.88
		SVM	0.89	0.89	0.89
		RF	0.88	0.88	0.88
		Ada Boost	0.88	0.86	0.88
		XGB	0.90	0.89	0.90
3	**ResNet-50**	KNN	0.89	0.87	0.89
		SVM	0.93	0.92	0.93
		RF	0.90	0.89	0.89
		Ada Boost	0.90	0.89	0.9
		XGB	0.92	0.92	0.91
4	**ResNet-152**	KNN	0.88	0.86	0.88
		SVM	0.93	0.92	0.91
		RF	0.98	0.97	0.98
		Ada Boost	0.90	0.90	0.91
		XGB	0.95	0.94	0.94

Table 3.3: Evaluation Measures on the Test dataset (Model 1)

SNo.	DCNN Model	Classifier	Accuracy	Precision	Recall (Sensitivity)
1	VGG-16	KNN	0.88	0.87	0.88
		SVM	0.91	0.90	0.91
		RF	0.89	0.88	0.89
		Ada Boost	0.88	0.87	0.88
		XGB	0.91	0.91	0.9
2	VGG-19	KNN	0.88	0.86	0.88
		SVM	0.90	0.89	0.9
		RF	0.88	0.87	0.88
		Ada Boost	0.89	0.87	0.89
		XGB	0.90	0.89	0.9
3	ResNet-50	KNN	0.88	0.86	0.88
		SVM	0.93	0.93	0.93
		RF	0.90	0.89	0.89
		Ada Boost	0.90	0.89	0.9
		XGB	0.92	0.92	0.92
4	ResNet-152	KNN	0.88	0.86	0.88
		SVM	0.93	0.92	0.91
		RF	0.98	0.97	0.98
		Ada Boost	0.90	0.90	0.91
		XGB	0.95	0.94	0.94

3.6 Model 2: Deep Feature Extraction with Neural Net Classifier

3.6.1 Proposed Architecture

In this approach, the DCNNs are deployed to extract features as well as for their further categorization. We utilized the following DCNNs: Mobile Net, VGG-16, VGG-19, Res-Net 50, Res-Net 152, and Dense-Net 169. The details of these DNNs are discussed in section 3.4. Figure 3.9 shows the design for the presented Model 2[113].

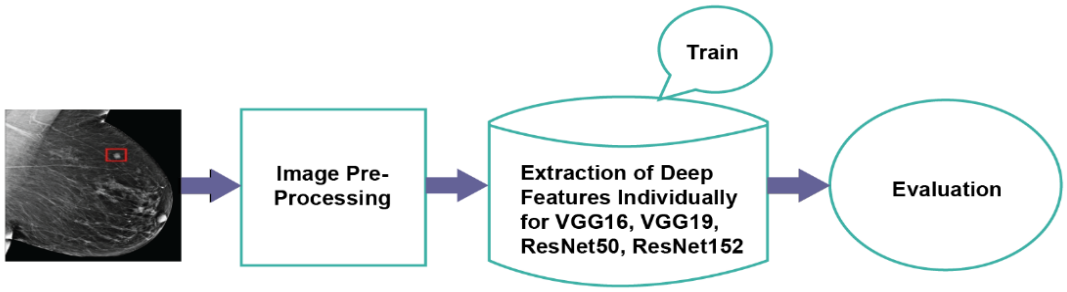

Fig. 3.9: Deep Feature Extraction with Neural Net Classifier

3.6.2 Evaluation Metrics

This segment discusses the performance measure for Model 2 i.e. Deep feature Extraction with Neural Net classifier. The following metrics have been assessed: Accuracy, AUC, Precision Recall, and Loss. Tables 3.4 to 3.7 show the corresponding measures on the Training, Validation, and Test dataset. These outcomes are analyzed in section 3.7.

Table 3.4: Training Metrics(Model 2)

S No.	DCNN Model	Accuracy	AUC	Precision	Recall	Loss
1	**Mobile Net**	0.99	0.99	0.96	0.95	0.03
2	**ResNet-50**	0.98	0.99	0.94	0.92	0.06
3	**VGG-16**	0.94	0.96	0.92	0.91	0.14
4	**VGG-19**	0.93	0.95	0.92	0.93	0.16
5	**DenseNet-169**	0.97	0.99	0.94	0.92	0.07
6	**ResNet-152**	0.96	0.98	0.92	0.90	0.08

Table 3.5: Validation Metrics(Model 2)

S No.	DCNN Model	Accuracy	AUC	Precision	Recall	Loss
1	**Mobile Net**	0.97	0.98	0.94	0.95	0.04
2	**ResNet-50**	0.95	0.96	0.94	0.95	0.07
3	**VGG-16**	0.92	0.93	0.90	0.91	0.08
4	**VGG-19**	0.90	0.92	0.89	0.90	0.06
5	**DenseNet-169**	0.92	0.94	0.92	0.91	0.06
6	**ResNet-152**	0.93	0.93	0.93	0.92	0.08

Table 3.6: Test Metrics(Model 2)

S No.	DCNN Model	Accuracy	AUC	Precision	Recall	Loss
1	**Mobile Net**	0.98	0.98	0.97	0.96	0.10
2	**ResNet-50**	0.96	0.95	0.95	0.93	0.16
3	**VGG-16**	0.92	0.95	0.91	0.92	0.17
4	**VGG-19**	0.91	0.94	0.90	0.89	0.19
5	**DenseNet-169**	0.93	0.95	0.93	0.91	0.17
6	**ResNet-152**	0.92	0.95	0.94	0.92	0.23

3.7 Experimental Evaluation & Performance Analysis

Considering the two suggested models, this section examines the visualizations and analysis of the Transnet framework. While the analysis of both models is covered in Section 3.7.2, the plots are shown in Section 3.7.1.

3.7.1 Plots of DCNN concerning Model 1 and Model 2

Figures 3.10 to 3.12 exhibit the visualizations of the Training, Validation, and Test dataset for Model 1(Deep Feature Extraction with ML Classifier).

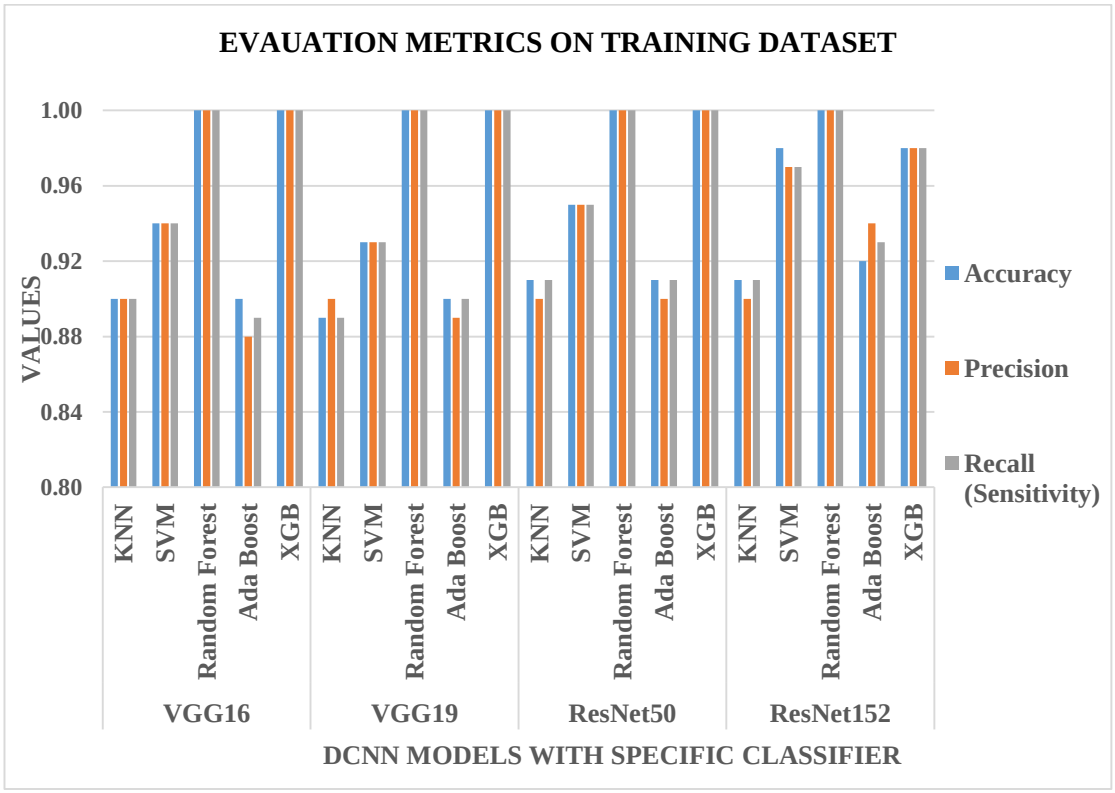

Fig. 3.10: Plots of DCNN models with various ML classifiers on the Training Dataset (Model 1)

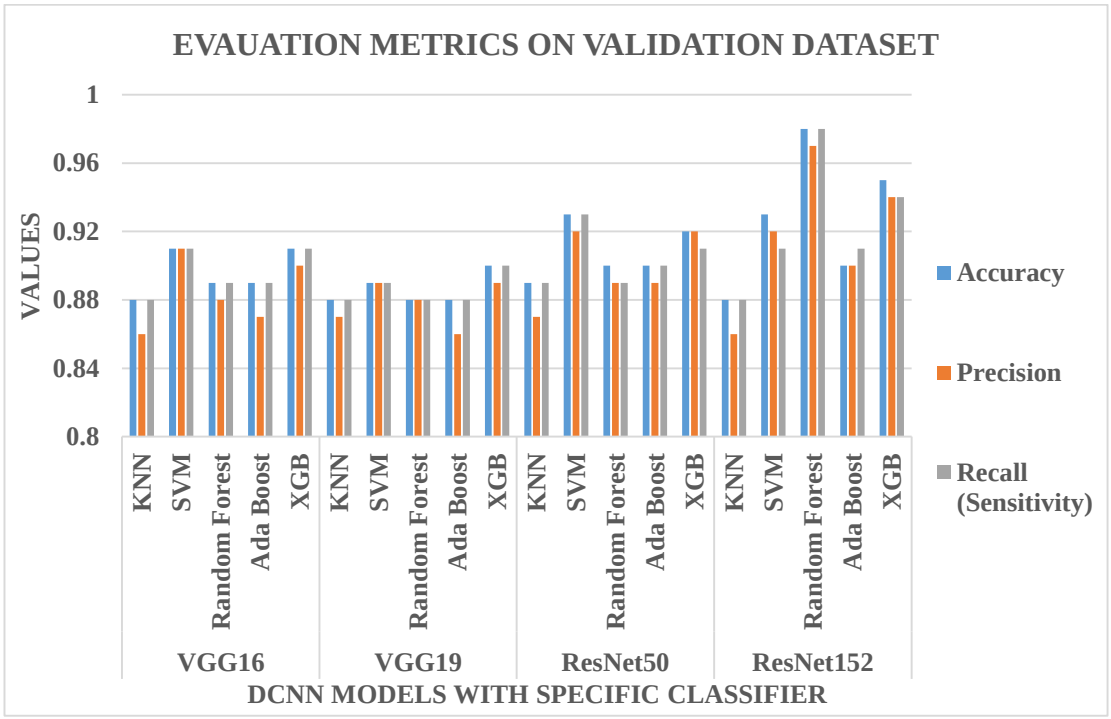

Fig. 3.11: Plots of DCNN models with different ML classifiers on the Validation Dataset (Model 1)

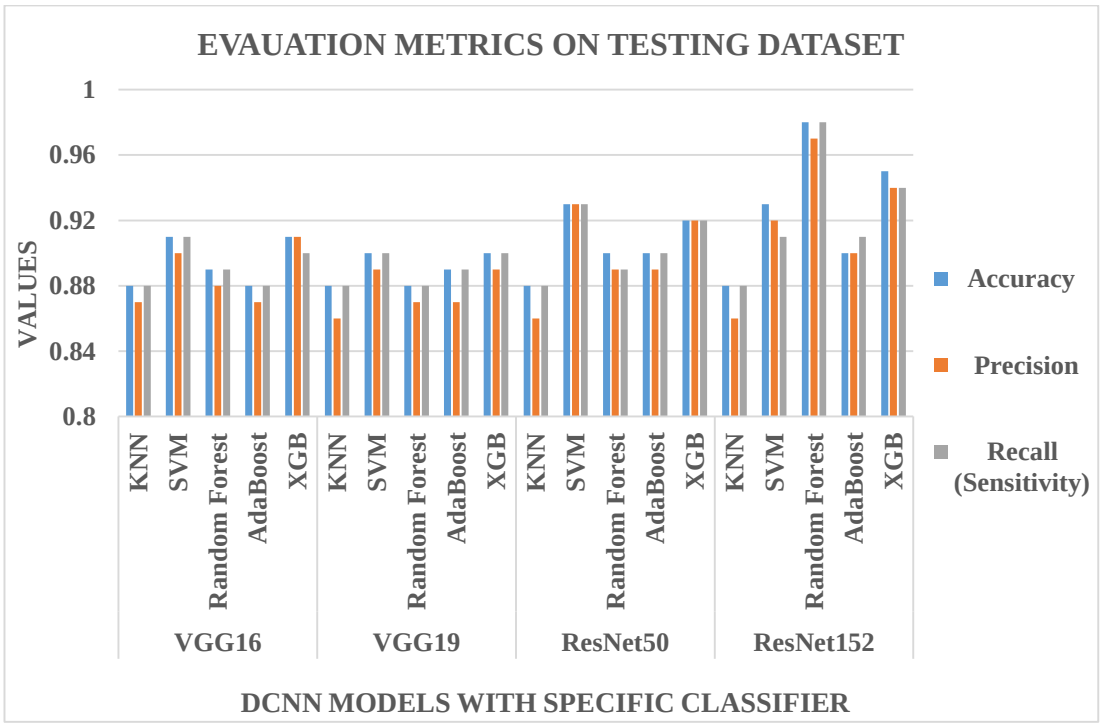

Fig. 3.12: Plots of DCNN models with several ML classifiers on the Testing Dataset(Model 1)

Figures 3.13 to 3.15 exhibits the Accuracy, AUC and Loss plots on the Training and Validation datasets for the proposed Model 2(Deep Feature Extraction with Neural Net Classifier).

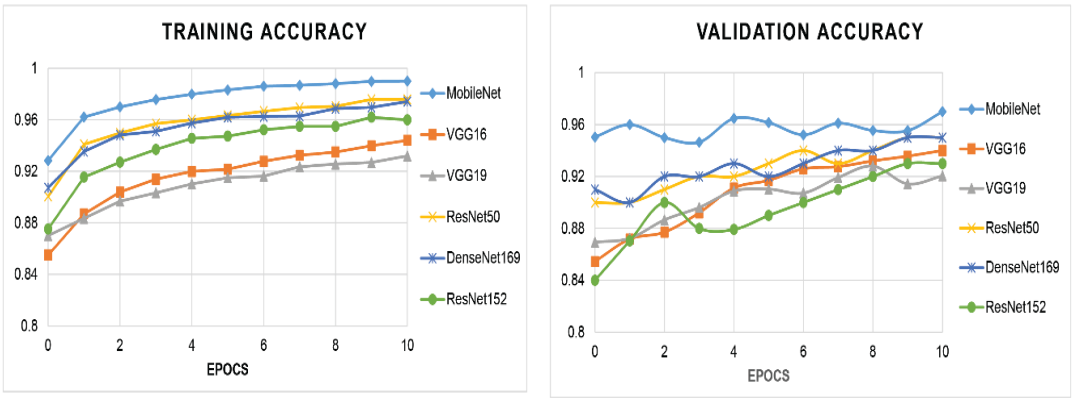

Fig. 3.13: Plots for Training and Validation Accuracies (Model 2)

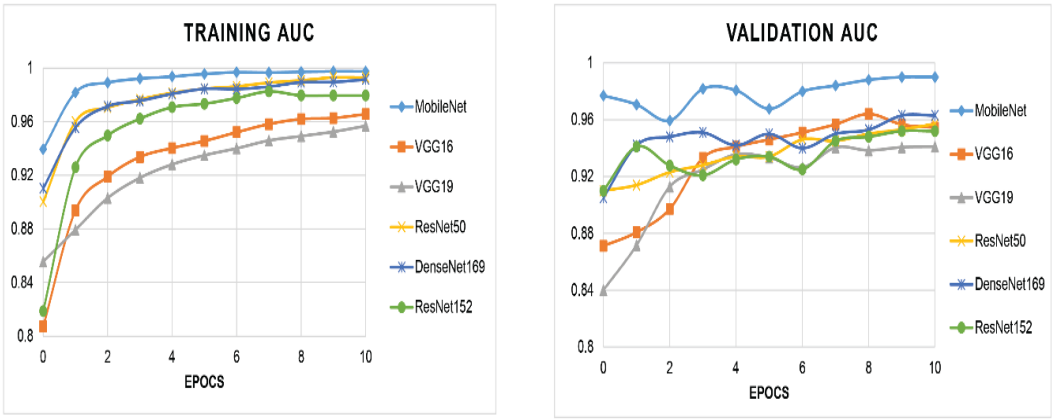

Fig. 3.14: Plots for Training and Validation AUC (Model 2)

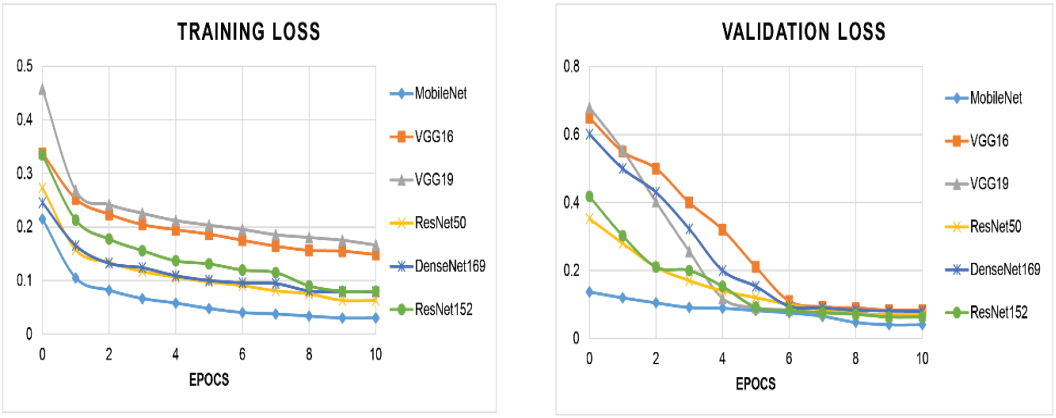

Fig. 3.15: Plots for Training and Validation Loss (Model 2)

3.7.2 Analysis: Model 1 and Model 2

The implementation was completed on the NVIDIA Tesla GPU system with 32 GB RAM. The framework utilized early stopping criteria and was initially trained for 20 epochs, which was set to minimize the validation loss with an interval of 5 epochs. The best weights during the training were restored at the end. The model underwent its initial training with a learning rate and batch size of .01 & 64, but the results were unsatisfactory. Thus, some hyper-parameters were optimized, and finally, the learning rate of .0001 and batch size of 128 were adopted. Both approaches utilize DNNs such as VGG-16, VGG-19, ResNet-50, ResNet-152, DenseNet-169, and Mobile Net v2. In the second approach (NN classifier), fine-tuning was performed for each model whereas in the first approach (ML Classifier), no fine-tuning was performed, and the raw outputs without any activation functions were captured as the features.

We compare the two proposed methods by evaluating the dataset on an 80:10:10 stratified split ratio for training, validation, and testing. 5 fold stratified cross-validation was executed and the mean of all observations across the folds was taken. The first approach, i.e., Deep Learning feature extraction with ML classifier uses pre-trained models with ImageNet weights. No training or fine-tuning was performed on these models, and the raw outputs without any activation functions were captured as the attributes. The derived attributes were supplied to the ML classifiers for binary

categorization. The ML classifiers varied from simplistic models like k-Nearest-Neighbour to sophisticated gradient-boosting models like XG-Boost. As evident from Table 3.1, during the training phase, Random Forest and XG-Boost exhibited 100% accuracy, precision, and recall rate on all pre-trained models, whereas simpler models like KNN could achieve an average of 90% (+- 2%) accuracy, precision and recall with a margin of 10% lower than other ML Classifiers. Support Vector Machine Classifier achieved intermediate results of 95% (+- 2%) accuracy, precision, and recall rate. These results closely matched the testing data, as demonstrated by Table 3.3. Thus, it was concluded that the deeper pre-trained networks like ResNet-152 outperform shallower models like ResNet-50, VGG-19, and VGG-16 by a margin of 6% increase in accuracy, precision, and recall rates. Additionally, Random Forest and SVM outperformed the other models by a margin of 5% in all observable metrics. This approach is less computationally complex as the pre-trained networks were not retrained.

The second approach, i.e., Deep Learning feature extraction with Neural Network Classifier, uses these models to execute feature extraction and classification. As a result, the computational complexity increased several fold. Dedicated GPU rigs were required to fine-tune the models. For the training phase, Mobile Net, ResNet-50, and DenseNet-169 performed better than older models like VGG-16 and VGG-19. The same metrics were observed during the test phase as well. The other models, however, closely matched the best-performing models. Their less complex nature may also be preferred during the deployment phase for faster predictions and lower memory footprint. Comparing the second approach against the first one, DCNN with Neural Net Classifier highlighted a 4% increase in performance by the classical machine learning models when paired with deep feature extraction techniques. This is because in this approach DCNNs were utilized for feature extraction and classification compared to ML classifiers in the first approach. In the first approach, no fine-tuning was performed whereas this approach fine-tuned the pertained networks as per the requirement of the target domain and thus increased the model's efficiency with high accuracy values. Hyper-parameter optimization was also performed which further contributes to enhancement in accuracy in the second approach.

3.8 Comparison of the Developed Model with Revolutionary Approaches

In this section, the proposed CAD model i.e. Transnet is compared with other reference techniques in the domain. The articles were compared from 2018 to 2023. As observed from Table 3.7, the Transnet model outperforms other cutting-edge techniques for Breast Carcinoma Classification.

Table 3.7: Comparative Analysis of Transnet Framework with Recent Approaches for Breast Carcinoma Diagnosis

Year	Reference	Imaging Modality	Feature extraction/ DNN Model Adopted	Dataset	Accuracy	AUC
2018	[133]	Mammography	CNN	MIAS	85.85	--
2018	[84]	Histopathology Images	CNN, LSTM Classification: SVM, Softmax	Break-His	91.00	
2018	[134]	Mammography	Generative Adversarial Network (GAN) +Res-Net	DDSM	--	0.89
2019	[135]	Mammography	DCNN: Alex Net Classification: SVM	CBIS-DDSM	87.2	0.94
2019	[136]	Mammography	Deep feature fusion of VGG-16, VGG-19, Google Net, and Res Net 50	CBIS-DDSM	96.6	0.93
2020	[137]	Mammography	Google Net Classification: XG Boost	DDSM	92.8	--

2020	[138]	Mammography	Feature fusion of several Models Classification: SVM, XG-Boost, Naïve Bayes, KNN, DT, Ada Boosting	CBIS-DDSM	90.91	--
2021	[38]	Mammography	Deep feature fusion of Alex Net, Google Net, ResNet-18, ResNet -50, ResNet101 Classification: SVM	MIAS	97.4	1.00
2021	[139]	Histopathology	VGG-19, ResNet-34, ResNet-50 along with Structural Pruning	Break-His	92.07: ResNet-50	--
2022	[140]	Mammography	ResNet-50, NasNet-Mobile Network	MIAS	89.5: ResNet-50 & NasNet	--
2022	[141]	Mammography	CoroNet (Based on the Xception Net Model)	CBIS-DDSM	94.92 (4-class) 88.67 (2-class)	
2022	[142]	Histopathology Images	VGG-16, VGG-19, Inception-ReNetv2, DenseNet-121, and DenseNet-201	Private Dataset	92.64: DenseNet-121	--
2023	[143]	Mammography	ResNet-50 Classification: KNN, RF, SVM	Private Dataset	85: KNN	0.89
2023	[144]	Histopathology	K-Means for Segmentation and ResNet-18 for feature extraction, SVM for classification	Break His	92.6	--

2023	[145]	Histopathology	3D U Net Model	Private Dataset	97	--
2023	**Proposed CAD Model TransNet**	Mammography	First Approach: VGG-16, VGG-19, Res Net50, Res Net 152 ML Classifiers: KNN, SVM, Random Forest, Ada Boost, XGB	CBIS-DDSM	**Best Result:** Train Set: 100 **(Random Forest & XGB)** Test Set: 98 (**ResNet-152 with Random Forest Classifier**)	--
			Second Approach: Mobile-Net, VGG-16, VGG-19, Res Net-50, Res Net-152, DenseNet-169. The same Models were used for the Classification	CBIS-DDSM	**Best Result**: Train Set: 99 Test Set: 98 (**Mobile Net**)	**Best Result:** Train & Test Set: 0.99 & 0.98 (**Mobile Net**)

3.9 Chapter Summary

The Chapter presents Transnet framework for diagnosing and classifying breast carcinoma on the CBIS-DDSM dataset. VGG-16, VGG-19, Mobile Net, ResNet-50, ResNet-152, and DenseNet-169 pre-trained networks were utilized during training. Two experiments were performed: In the first approach, namely Deep feature fusion with ML Classifier, pre-trained networks are deployed as feature extractors, and afterwards the derived attributes are provided to ML classifiers for categorization. The second approach, called Deep feature fusion with Neural Net classifiers, fine-tune these networks for feature extraction and categorization. The chapter concludes with a comparative analysis of the presented approach with cutting-edge techniques. The findings reveal that the stated system performs superior to other cutting-edge

approaches. In the future, this framework could be developed on other imaging modalities.

PUBLICATION

The following journal publishes the research covered in this chapter:

- G. Chugh, S. Kumar, and N. Singh, “TransNet: a comparative study on breast carcinoma diagnosis with classical machine learning and transfer learning paradigm,” *Multimed. Tools Appl.*, no. 0123456789, 2023, doi: 10.1007/s11042-023-16938-x. **(SCIE Indexed, IF=3.0)**

CHAPTER 4

A MULTI-STAGE TRANSFER LEARNING PARADIGM FOR BREAST CARCINOMA DIAGNOSIS

4.1 Overview

DCNNs are among the optimal learning algorithms for analyzing pictures and have demonstrated outstanding performance in diagnosis. As discussed in Chapter 2, DCNN could be trained either from the beginning or by a transfer learning approach. DCNNs trained from scratch don't utilize any pre-trained architecture and consume many resources during training. They require high computing GPUs in the training stages. The transfer learning approach is employed when the training data from the target destination is significantly less; thus, the expertise acquired from the original discipline is transferred to the target discipline. The Fine-Tuning strategy is adopted, and therefore the network is fine-tuned layer-wise. Fine-tuning involves adapting the last layers of the network as per requirements. This chapter proposes the Multi Stage Transfer Learning Approach(MSTLA) for diagnosing and categorizing breast malignancy. DenseNet-169 and ResNet-152 are utilized with a three-stage transfer learning strategy. The results show that both DenseNet-169 and ResNet-152 have outstanding performance in the third stage of transfer learning.

4.2 Dataset Utilized

In this approach, we have utilized three mammogram datasets - MIAS, In-Breast, and DDSM. The detailed description of the mammogram scanners and machines used could be referred to in[42], [46], [112]. A brief overview of these datasets is discussed below:

(i) **CBIS-DDSM**[112]**:** It is the largest dataset for mammograms, comprises of 2620 cases with 10,480 images. It contains both CC and MLO views of mammograms. The patient's age, family history, and Breast Imaging Reporting and Database System (BI-RADS) score are also incorporated in the dataset directory.

(ii) **In-Breast**[46]**:** In-Breast is another popular dataset of mammograms. It comprises 115 cases with 410 images. Data related to the patient's age, family history, and the BI-RADS score is provided in the dataset.

(iii) **MIAS**[42]**:** Mammography Image Analysis Society(MIAS) is another dataset for mammograms. It comprises 161 cases having 322 mammogram images with a Medio lateral Oblique (MLO) view. The instances are classified into normal, benign, and malignant.

4.3 Deep Neural Networks (DNNs)

In the proposed MSTLA framework two DNNs i.e. DenseNet-169 and ResNet-152, are fine-tuned. The detailed architecture of both networks is discussed in Chapter 3. An overview of both networks are discussed below:

(a.) DenseNet[121]**:** Dense Nets are densely connected convolution neural networks. Dense Net joins the outcome of the former layer with the succeeding layer. It was designed to enhance the loss in accuracy produced by the vanishing gradient in deep neural networks. It also reduces feature count and permits feature reuse. The proposed work fine-tunes the Dense - Net-169 model with 169 layers of depth. To obtain the optimized results, the model has been fine-tuned layer-wise in three stages which is further elaborated in Section 4.4.2.

(b.) ResNet[116]**:** Res-Net, also called Residual Network proposed the notion of Residual Blocks to overcome the issues caused by vanishing gradients. These

networks use a strategy called Skip Connection that allows layers to connect and skips some layers in between. ResNet-152 architecture with 152 layers of depth has been utilized in the work and has been fine-tuned layer-wise separately in three stages. Epochs defined at the first stage were carried forward to the consecutive stages. Section 4.4.2 elucidates the proposed architecture with further details of the proposed stage-wise fine-tuning.

4.4 Formulated Model for Multistage Transfer Learning

4.4.1 Pre-Processing & Data Augmentation

Pre-processing of images is a process of cleaning and enhancing the images to convert them into the form fed to a DNN. The steps we have followed to pre-process the datasets include Min-Max Normalization, CLAHE (Contrast Limited Adaptive Histogram Equalization), and Padding. These steps are discussed in detail in Chapter 3.

During the training stage, DCNN requires a substantial volume of data. Insufficient training data may result in issues like overfitting, where the network works effectively with training data but performance worsens on testing data. The images in medical datasets are not available in abundance. The technique of "data augmentation" involves increasing the dataset's size. Thus, a series of operations, including Image Rotation and flipping, are performed to enhance the dataset size. Following data augmentation on the three datasets, the MIAS dataset contains 3,816 images (Benign-2376 & Malignant-1440), the In-Breast dataset has 7,632 images (Benign-2520 &Malignant-5112) and the CBIS-DDSM dataset encompasses 13,128 images (Benign-5970 & Malignant-7158).

A brief description of Pre-Processing and Augmentation is depicted in Figure 4.1.

4.4.2 Proposed Architecture

In this chapter, we are utilizing the concept of MSTLA wherein multiple stages are used for training the model through a transfer learning approach[146]. The steps taken in the procedure are listed below:

(i) Stage 1: Training on the MIAS dataset

This is the first step in MSTLA. In Stage 1, pre-processed mammograms from the MIAS dataset were trained on DenseNet-169 and ResNet-152 models. Initial layers were kept frozen, and the first fine-tuning was applied at the 100^{th} layer.

(ii) Stage 2: Training on In-Breast dataset

In this stage, the resultant DCNN obtained from stage 1 is further trained on the pre-processed images from the In-Breast dataset. Fine Tuning was performed at the 350^{th} layer, and the subsequent DCNN obtained is forwarded to the next stage.

(iii) Stage 3: Training on the DDSM dataset

This is the last stage, here we utilized pre-processed images from the CBIS-DDSM dataset for the training of DCNN obtained from Stage 2. The last fine-tuning was applied at the 450^{th} layer. The results obtained from each stage were evaluated and compared.

The architectural diagram for the stated framework is illustrated in Figure 4.1.

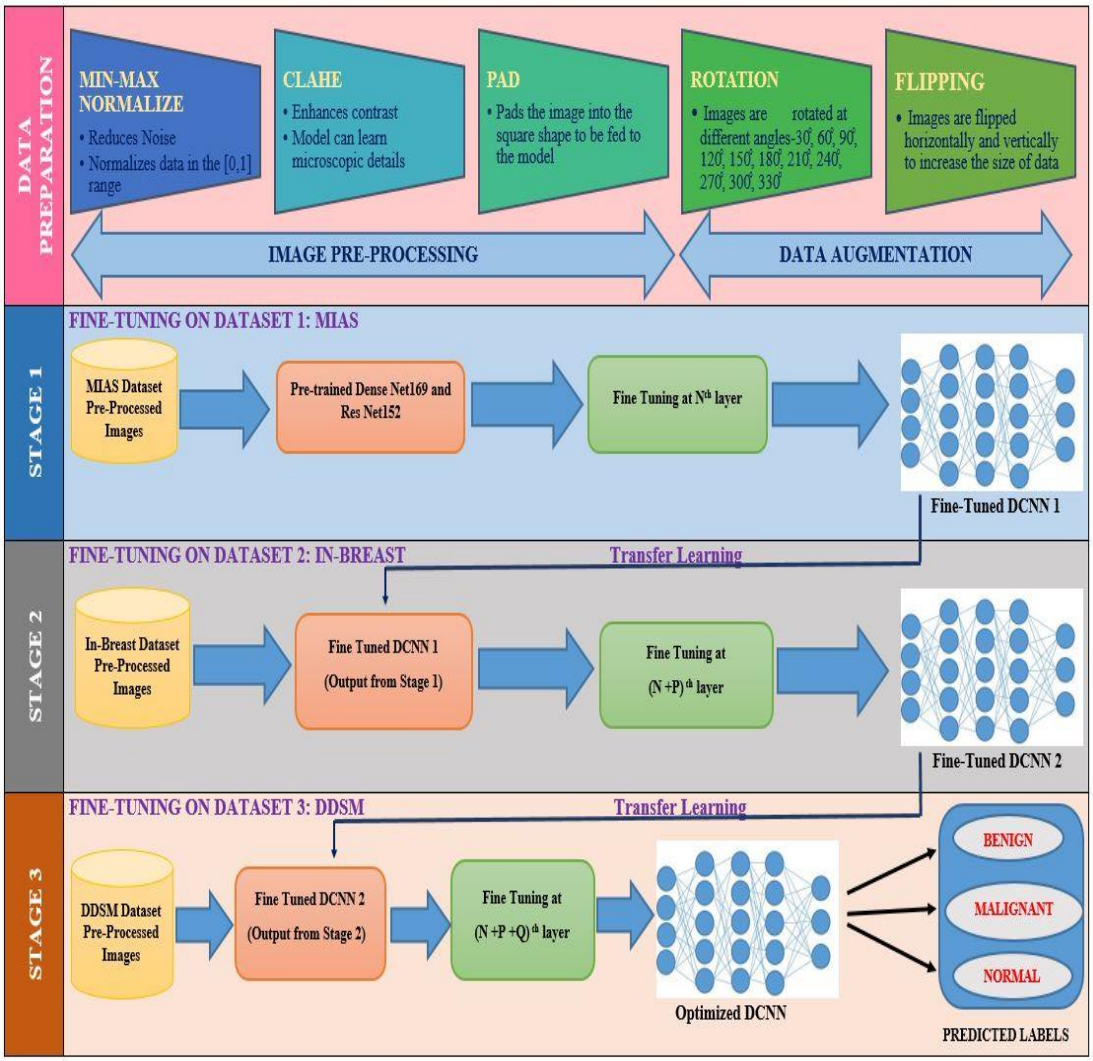

Fig. 4.1: Proposed MSTLA Methodology

4.4.3 Evaluation Metrics

The stated framework has been assessed with the following metrics: Accuracy, AUC, Precision, Recall, and Loss. Cross Entropy Loss i.e. Log Loss has been examined here (Discussed in Chapter 2). Tables 4.1 and 4.2 show the corresponding Results on Training and Validation datasets[146].

Table 4.1: Training Set Evaluation Metrics

Training Set	Models	Accuracy	AUC	Precision	Recall	Loss
Stage 1	**DenseNet-169**	0.98	0.99	0.98	0.98	0.12
	ResNet-152	0.95	0.98	0.95	0.95	0.23
Stage 2	**DenseNet-169**	0.99	0.99	0.99	0.99	0.08
	ResNet-152	0.97	0.99	0.97	0.97	0.11
Stage 3	**DenseNet-169**	0.99	0.99	0.99	0.99	0.002
	ResNet-152	0.99	0.99	0.99	0.99	0.005

Table 4.2: Validation Set Evaluation Metrics

Validation Set	Models	Accuracy	AUC	Precision	Recall	Loss
Stage 1	**DenseNet-169**	0.65	0.73	0.65	0.65	0.32
	ResNet-152	0.63	0.67	0.63	0.63	0.45
Stage 2	**DenseNet-169**	0.77	0.79	0.77	0.77	0.20
	ResNet-152	0.70	0.74	0.70	0.70	0.33
Stage 3	**DenseNet-169**	1.00	1.00	1.00	1.00	0.00008

	ResNet-152	0.99	0.99	0.99	0.99	0.01

4.5 Experimental Evaluation & Performance Analysis

This section examines the implementation details, visualizations, and analysis of the MSTLA framework. The charts are depicted in Section 4.5.1 and the implementation details along with the results analysis are discussed in Section 4.5.2.

4.5.1 Visualizations

This section explores the Accuracy, AUC, and Loss Plots on the Training and Validation Data Sets. The visualizations are shown stage-wise. We can observe from Figures 4.2 and 4.3, the Loss plots are not visible for Stage 3 as this stage depicts minimum loss.

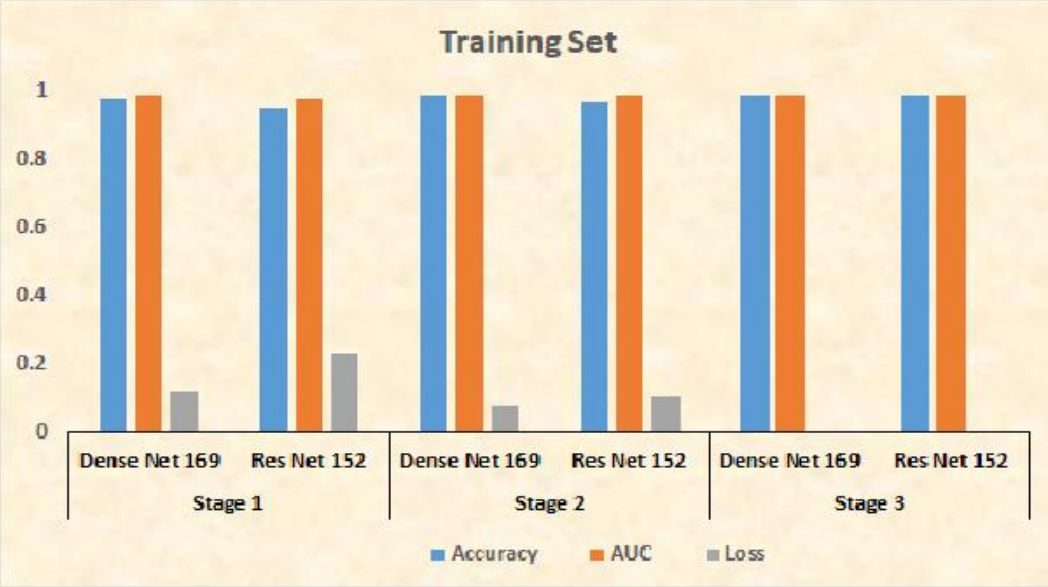

Fig. 4.2: Accuracy, AUC, and Loss Plots on Training Data Set

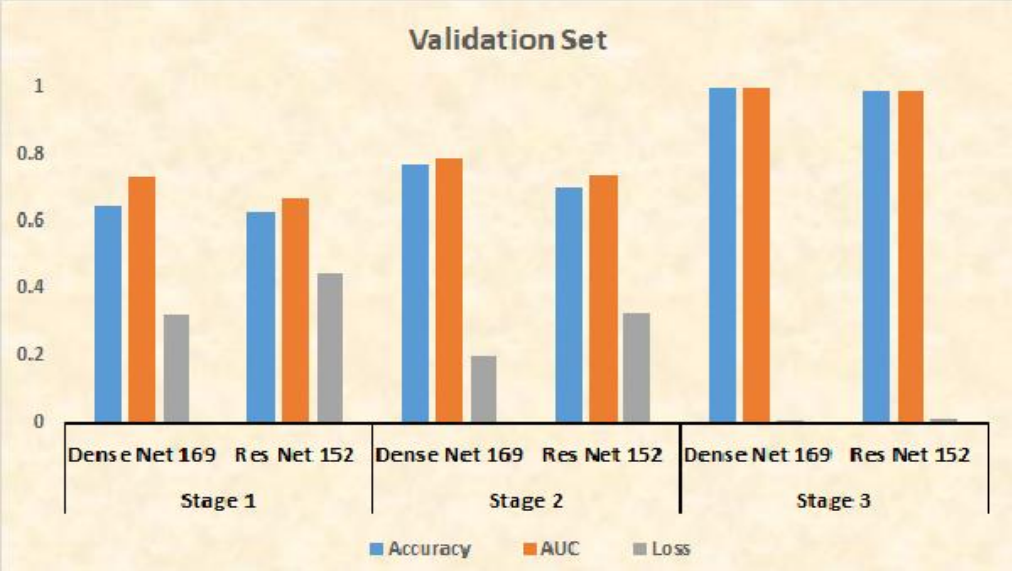

Fig. 4.3: Accuracy, AUC, and Loss Plots on Validation Data Set

4.5.2 Performance Analysis

We have designed a framework with a three-stage transfer learning approach and utilized three datasets- MIAS, In-Breast, and DDSM. Two popular DCNNs, i.e., DenseNet-169 and ResNet-152, were employed during training. The experiments were performed on the NVIDIA Tesla system with 32 GB RAM GPU. The epochs were defined at each stage, and the epochs from the previous stage were carried further in the next stage. Stage 1 began the training with ten epochs, followed by Stage 2 with 25 epochs, and subsequently followed by Stage 3 with 25 epochs. The early stopping criteria were initialized for each stage to optimize the validation loss with a patient interval of 5 epochs. Adam optimizer was utilized during the network's training with a learning rate and batch size of .0001 and 32, respectively.

Tables 4.1 & 4.2 show the performance metrics obtained after applying the stated approach on DenseNet-169 and ResNet-152. Both the networks performed exceptionally well in the proposed method and have shown remarkable performance for breast carcinoma diagnosis. During the training phase, DenseNet-169 and ResNet-152 performed very well with accuracy and an AUC value of 97% +-2 and 0.97+-0.2 in all three stages. The loss was a little high in Stage 1, started declining from Stage 1, and was minimal in the third stage, as apparent from Table 4.1.

During the validation phase, when the network was first trained with MIAS mammograms with fine-tuning at the specified layer, the accuracy and other parameters were relatively low. Table 4.2 shows that during Stage 1, the values of all the parameters were in the range of 0.65+- 0.2, and the loss was high. When the optimized DNN was further fine-tuned with the In-Breast data set during Stage 2, the values improved, and Stage 2 exhibited values in the range of 0.77 +-0.3. The loss started declining, but the decline was low. The last stage, i.e., Stage 3, fine-tuned the model further, and as a result, the model showed remarkable performance at this stage. DenseNet-169 has shown 100% accuracy and AUC with a minimum loss of 0.08e-3, whereas ResNet-152 exhibited 99.6% accuracy and AUC with a loss of .01.

4.6 Comparative Analysis of Developed Framework with Cutting-Edge Techniques

Table 4.3 compares the presented approach with other contemporary strategies employing the transfer learning technique. As evident from Table 4.3, the proposed model has shown remarkable performance for breast cancer diagnosis and outperforms the other cutting-edge technologies in the domain.

Table 4.3: Comparative Evaluation of MSTLA Framework with Cutting-Edge Techniques

Reference	Image Dataset	Approach	Model	Accuracy	AUC
[102]	Mammograms and Digital Breast Tomosynthesis Images (Private Dataset)	Transfer learning (Two-Stage)	Deep CNN	Not Evaluated	0.91
[62]	Mammograms (DDSM)	Transfer learning (One-Stage)	ROI-based CNN-You Look Only Once(YOLO)	97.0	0.96

[135]	Mammography (DDSM)	Transfer learning (One-Stage)	Alex Net	87.2	0.94
Proposed Approach	**Mammogram (MIAS, IN-BREAST, DDSM)**	**Multi-Stage Transfer Learning (Three Stage Approach)**	**Dense Net-169 and Res Net-152**	**Dense Net-169: 100** **Res Net-152: 99.6**	**Dense Net-169: 1.00** **Res Net-152: 0.99**

4.7 Chapter Summary

In this chapter, we have designed a model for diagnosing breast carcinoma using a Multi-Stage Transfer Learning Approach (MSTLA). We have used three datasets: MIAS, In-Breast, and DDSM. The model is fine-tuned in three stages on separate datasets, and the optimized DCNN is carried forward at the next stage. Two DCNNs are deployed for training the model – DenseNet-169 and ResNet-152. The results have shown that, in the training phase, both the DCNNs performed exceptionally well in all three stages; however, the results were quite different in the case of the validation dataset. In the validation phase, Stage 3 performs best compared to the other two stages, with DenseNet-169 having accuracy and AUC values of 100 and 1.0. On the other hand, ResNet-152 exhibits accuracy and AUC of 99% and 0.99. The results have shown that the proposed methodology could be employed for early-stage breast carcinoma diagnosis. In the future, the proposed methodology could be expanded to other modalities.

PUBLICATION

The research addressed in this Chapter has been published in the following IEEE Conference:

- G. Chugh, S. Kumar, and N. Singh, “MSTLA: Multi-Stage Transfer Learning Approach for Breast Carcinoma Diagnosis,” *2023 Int. Conf. Adv. Comput. Comput. Technol. InCACCT 2023*, pp. 509–514, 2023, doi: 10.1109/InCACCT57535.2023.10141697.

CHAPTER 5

GENERALIZATION ERROR IN DEEP CONVOLUTIONAL NEURAL NETWORKS

5.1 Overview

Deep learning has introduced various paradigms in the healthcare industry. Deep Convolutional Neural Networks assist doctors in diagnosis, surgery, and other areas. Thus, it is necessary to diagnose it so that the mortality count can be decreased. Generalizability defines the effectiveness of the network on the unseen data. When capturing mammograms different types of noise get added to the images. Noise may significantly diminish classification ability and make class separation more difficult. Thus, analyzing the model's generalization on noisy or unseen data is very crucial. This chapter proposes an approach for analyzing generalization errors in Deep Convolutional Neural Networks by inducting noises such as Gaussian, Salt and pepper, and Speckle. We have utilized the CBIS-DDSM dataset. Three prominent deep neural network models- Inception v3, DenseNet-201, and EfficientNet-B4- were fine-tuned to assess the model's efficiency on noisy data and thus Generalization error was evaluated.

5.2 Data Set Employed

We have utilized the CBIS-DDSM dataset[112]. This consists of 2620 instances with 10,480 images and is one of the largest databases for mammography. Several augmentation approaches, such as rotations, flipping etc. are utilized to avoid overfitting and to increase dataset size[140]. After implementing these techniques, our dataset size increased to 13,128 images.

5.3 Deep Convolutional Neural Networks

The following DCNNs have been utilized to evaluate the generalization capability and thus evaluate Generalization error:

(i) **Inception v3**[147]: Inception Networks reduces the overall computational cost that was incurred in traditional networks by introducing the concept of networks within networks. It comprises inception blocks with varying levels of convolutional filters. 1x1 convolution layer reduces the input data dimensions and assists in recognizing the depth details of the network across all of the image's channels. Convolutions of 3 x 3 and 5 x 5 dimensions are used to learn spatial characteristics at various scales. Several variations in Inception modules were proposed that include: Inception v1, Inception v2, and Inception v3. In this work, we have utilized Inception v3 architecture. Figures 5.1, 5.2, and 5.3 visualize the layered diagram of Inception v3.

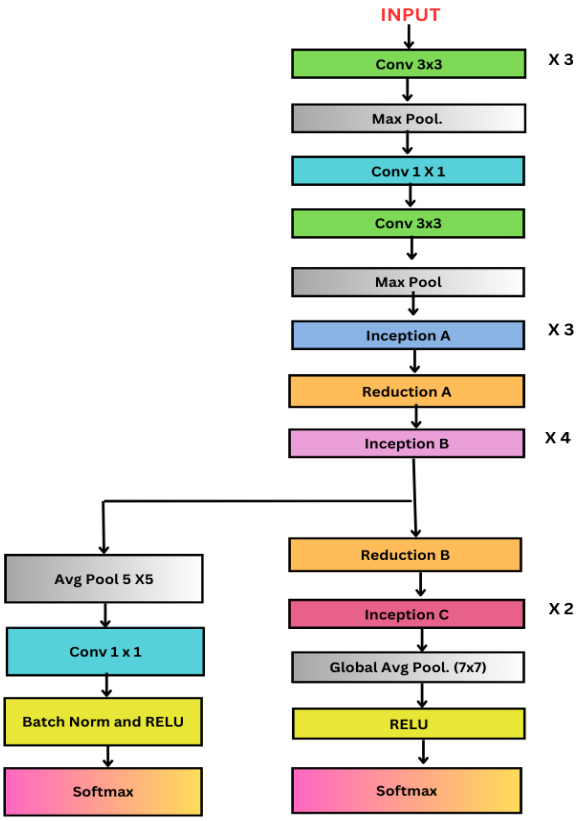

Fig. 5.1: Inception v3 Architecture[147], [148]

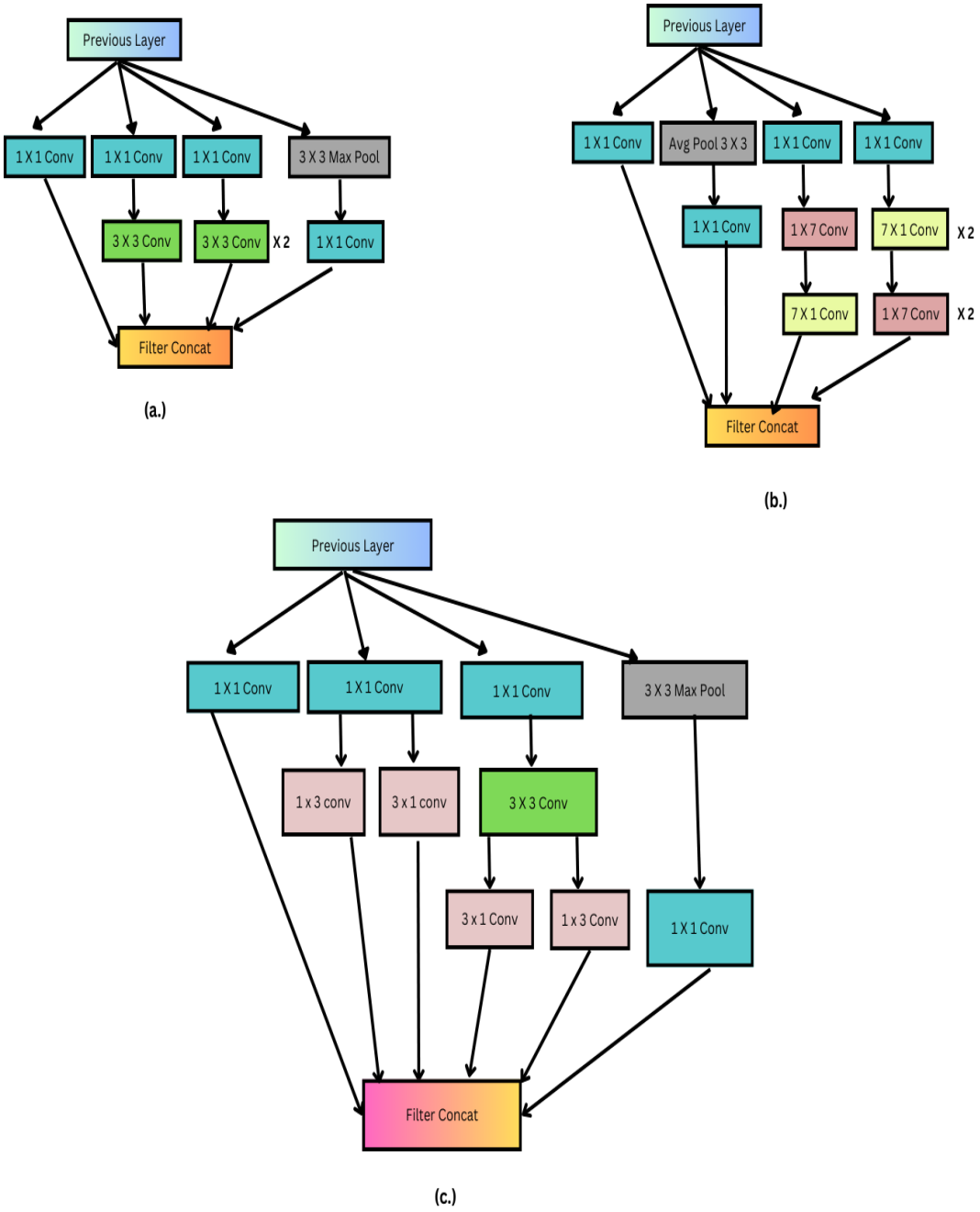

Fig. 5.2: Inception v3 Elements: (a.) Inception A (b.) Inception B (c.) Inception C[147], [148]

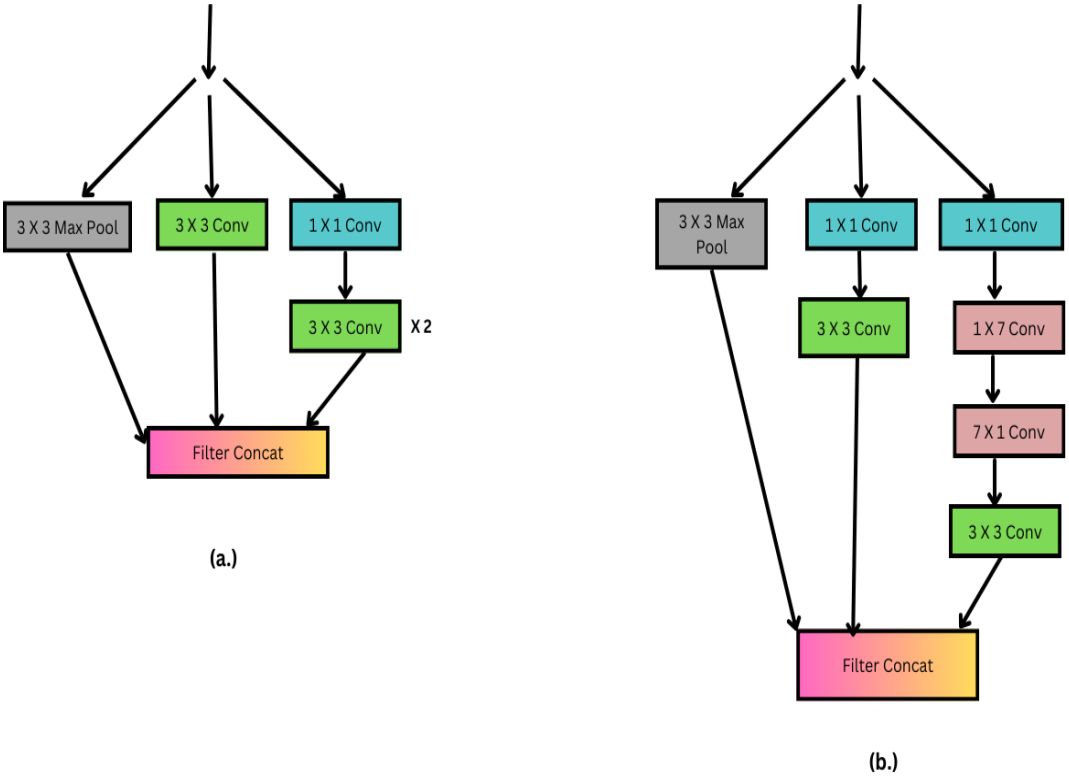

Fig. 5.3 Inception v3 Sections: (a.) Reduction A (b.) Reduction B[147], [148]

(ii) **DenseNet-201**[121]**:** Densely Populated Convolutional Networks have dense connectivity patterns and thus ensure maximum information flow. These networks connect each layer to all the succeeding layers and overcome the challenge of vanishing gradients in traditional networks. Dense Net reduces feature count and allows feature reuse. Feature maps learned from several layers are concatenated and thus efficiency could be improved. These networks come with varying levels of depths- DenseNet-121, DenseNet-161, DenseNet-169, DenseNet-201, etc. This work explores the DenseNet-201 model. The architecture of DenseNet-201 with 201 layers' depth is shown in Figure 5.4.

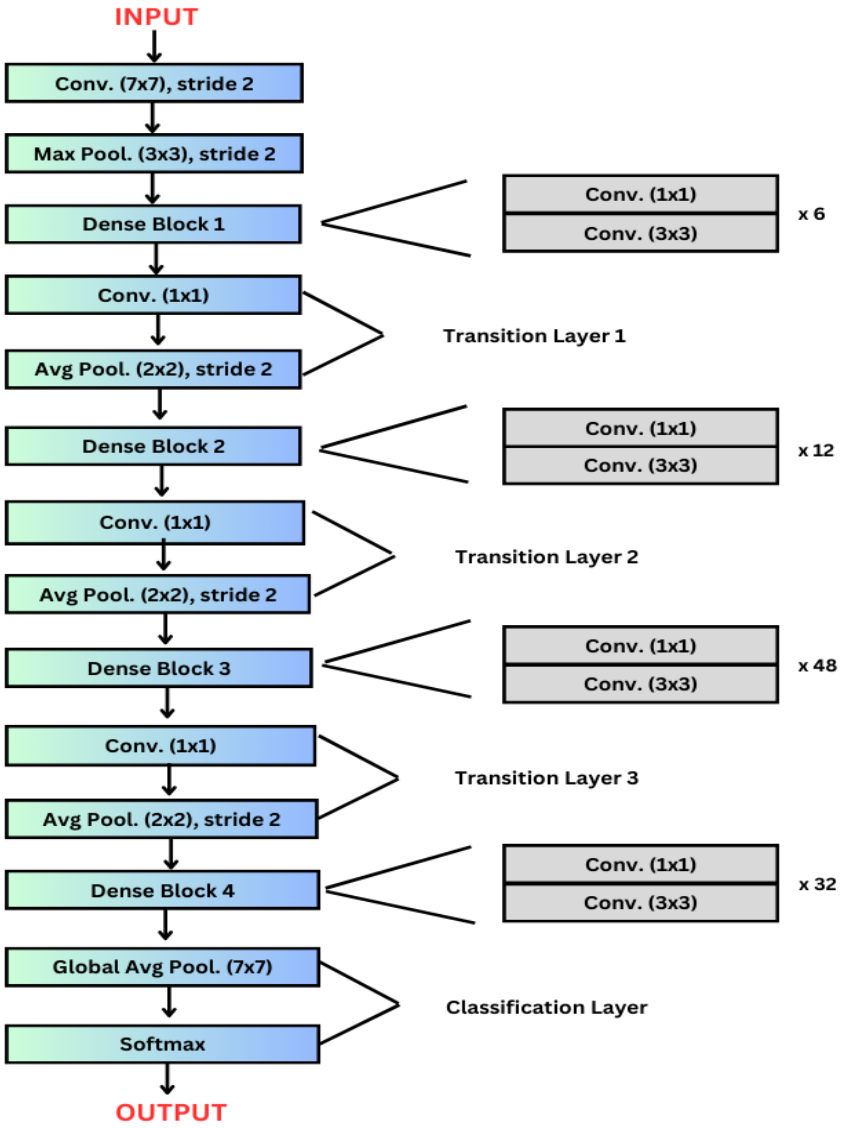

Fig. 5.4: DenseNet-201 Architecture[121]

(iii)**EfficientNet-B4**[149]**:** Convolutional Neural Networks are scaled up in various dimensions to achieve better accuracy and efficiency. The Efficient Net model utilizes the compound scaling method with fixed scaling coefficients for depth, width, and resolution. Mobile inverted bottleneck convolution (MB Conv) with squeeze and excitation optimisation is the fundamental building component of the Efficient Net architecture[150]. By varying several parameters, different versions of the model are obtained: EfficientNet- B0, EfficientNet-B1, EfficientNet-B2, EfficientNet-B3, EfficientNet-B4, Efficient

Net-B5, EfficientNet-B6, EfficientNet-B7. These scaled models efficiently reduce parameter count and Floating Point Operations per second (FLOPS). The efficient Net B4 version is exploited to assess the generalization error (Figure 5.5).

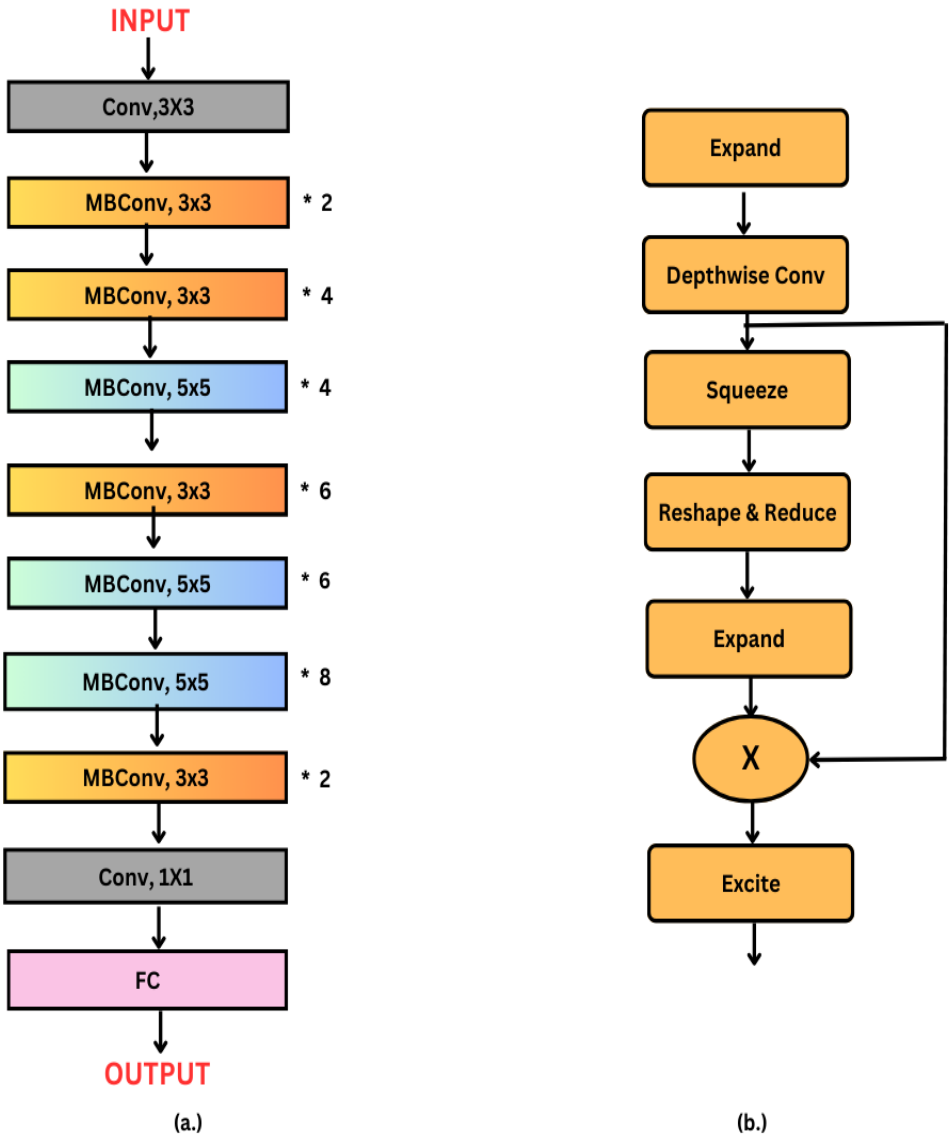

Fig. 5.5: EfficientNet-B4 Architecture[149]–[151]: (a) Layered Diagram (b.) Building Block: MB Conv

5.4 Formulated Model for Generalization Error Estimation

This section explores the basics of generalizability and generalization error in deep convolutional neural networks.

5.4.1 Introduction

Significant research has been performed on diagnosing breast carcinoma using machine learning and deep learning strategies [38], [113], [144]–[146]. These studies have proposed several methodologies for early diagnosis of Breast Carcinoma and have attained pretty good accuracies. Very little research in the survey emphasizes on the generalizability of DNNs. The inability of DNNs to generalize across domains is a significant barrier to their use in improving medical scenarios [106].

Due to their large learning capacity, Deep Convolutional Neural Networks(DCNN) can memorize the training data. Generalization is an approach to analyze how the model behaves on unseen data. Generalization Error(GE) measures the difference between training and testing errors. Leading causes of GE include memorization of training data, overfitting, a model with too many parameters, etc. When Deep Neural Networks are trained on medical imaging datasets, it is essential to analyze the effect of generalization error on transfer learning networks[105]. The performance of any model depends upon the generalization ability. Regularization approaches for example data augmentation, and dropout reduce the model complexity and thus improve the generalization capability of DNN by preventing overfitting and thus leading to low GE.

Further, the robustness of any DNN can also be examined through various Adversarial Attacks. Adversarial attacks represent any perturbations being added to the dataset to befool a deep learning model. In contrast to non-medical DL models, medical DL networks are more susceptible to adversarial attacks [152]. Thus, there is an urgent need to study and analyze the impact of these attacks on our models.

This chapter presents a novel technique to assess the Generalization error through various Noise samples. The noise affecting Mammogram images includes Gaussian, Salt & Pepper, and Speckle. We have analyzed and plotted the model behaviour

concerning the various perturbations added to the dataset to analyze the generalization capability of the DCNN.

5.4.2 Noise Categorization

Noise represents undesired information in digital images. During mammogram acquisition, a small perturbation can corrupt the entire image, and thus it's very crucial to examine the impact of noise to study the generalizability of models. The following categories of noise could be embedded in mammograms:

(i) **Gaussian Noise**

This is also referred to as Additive/Amplifier noise. It is additive in nature and corrupts the grey values in the images[153]. It utilizes standard Gaussian distribution and is given as:

$$PD(z) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(z-u)2}{2\sigma^2}} \quad (5.1)$$

Here, PD(z)= Probability Density Function for Gaussian Noise ; z=Grey Level; u=Mean of the average of z ; σ=Standard deviation; σ^2=Variance

(ii) Salt and Pepper Noise

This is referred to as Impulse noise and it worsens the overall quality of the image. The image shows it as a series of black and white dots. i.e. we see dark dots in bright areas and bright dots in dark places[153]. The probability density function is given as:

$$PD(z) = \begin{cases} P_{\mathrm{m}} & for\ z = m \\ P_{\mathrm{n}} & for\ z = n \\ 0 & otherwise \end{cases} \quad (5.2)$$

Here, PD(z)= Probability Density function for Impulse Noise; z= Pixel Value; m, n = grey level values;

if n>m= level m will appear as a darker spot and

if n<m= level m will appear as a lighter spot

if P_m= P_n=0, it is called Unipolar noise

(iii) Speckle Noise

Speckle noise comes under the category of multiplicative noise. This is also called granular noise. It diminishes the fine details and edge sharpness, leading to deterioration in image quality in a way similar to Gaussian noise. Speckle noise might have an impact on the segmentation and classification processes in the biomedical processing of images[154]. Speckle Noise can be represented as:

$$\mathrm{I_{noisy}(m,n) = I_{original}\ (m,n)\ *\ N_{speckle}(m,n)} \quad (5.3)$$

$$\mathrm{N_{speckle}(m,n) = 1 + \eta(m,n)} \quad (5.4)$$

In the above equations, (m, n) represents a particular pixel; $I_{noisy}(m, n)$ represents the intensity of the noisy image; $I_{original}\ (m, n)$ represents the intensity of the original image; $N_{speckle}(m, n)$ represents speckle noise; $\eta(m, n)$ represents Gaussian noise with zero mean and variance σ^2

5.4.3 Proposed Architecture

The framework to estimate generalization error in DCNNs is divided into the following two phases[104]:

(i) In Phase 1, we fine-tune tune the following DCNNs i.e. Inceptionv3, DenseNet-201, and EfficientNet-B4 on the CBIS-DDSM dataset. The metrics are evaluated on the trained and validated model.

(ii) In Phase 2, we analyze the generalizability of DCNNs by corrupting the dataset through the insertion of three types of noises. At the next step, we again fine-tune the following DCNNs i.e. Inceptionv3, DenseNet-201, and EfficientNet-B4 on the corrupted dataset. The results are computed on the corrupted dataset.

Generalization error is thus computed by calculating the difference in loss rates on the corrupted and validation datasets. The proposed structure to estimate generalization error is shown in Figure 5.2.

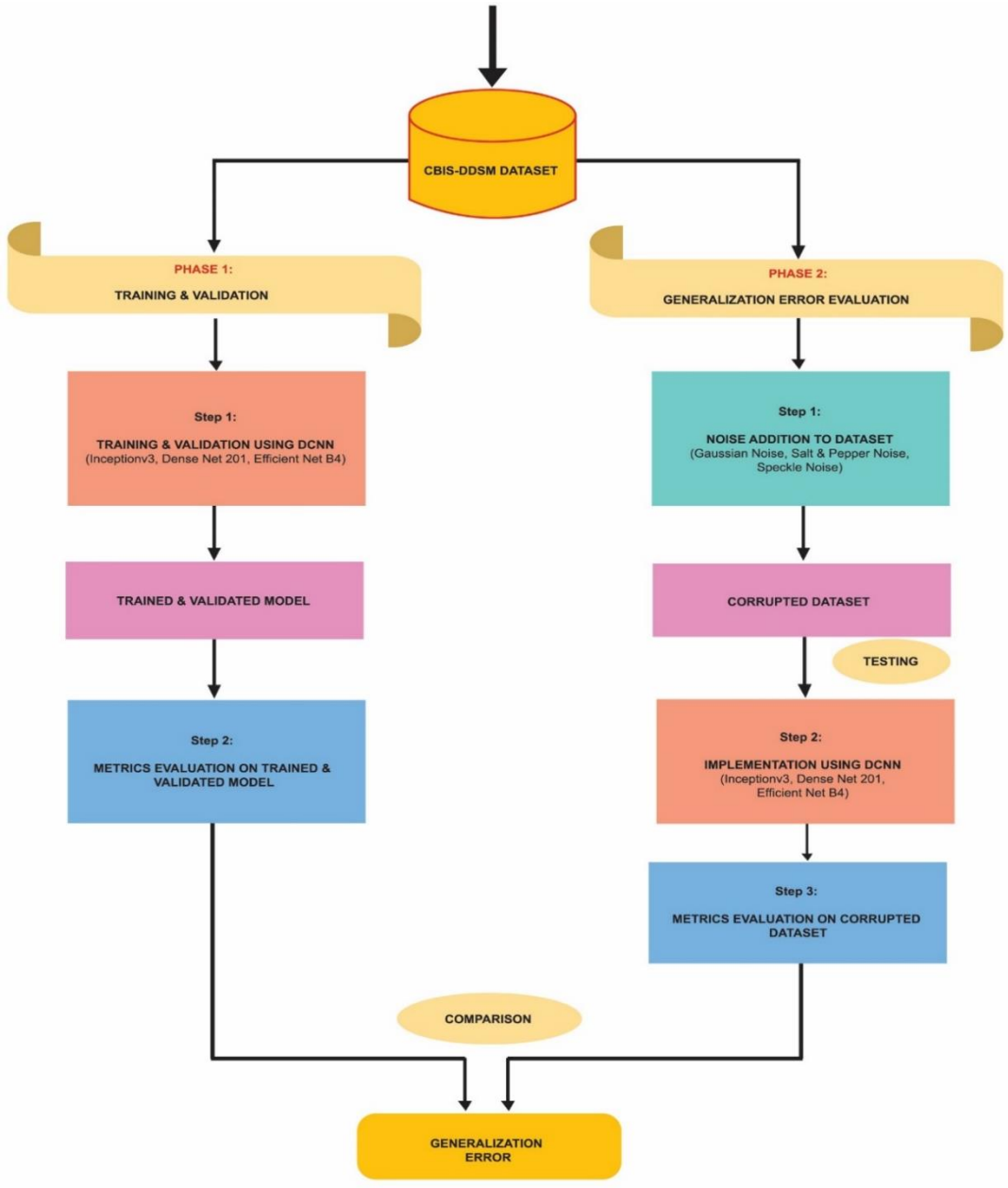

Fig. 5.6: Proposed Architecture for Generalization Error Evaluation

5.4.4 Evaluation Metrics

In this work we have evaluated following metrics: Accuracy, AUC, Precision, Recall and Loss on the validation and Corrupted dataset. The loss rates of corrupted and

validated dataset are compared and thus generalization error is evaluated. Tables 5.1 and 5.2 depict these metrics and Table 5.3 shows the computed Generalization error for the three DCNNs.

Table 5.1: Performance Measures on the Validation Dataset

S No.	DCNNs	Accuracy	AUC	Precision	Recall	Loss
1	**Inception v3**	0.98	0.99	0.98	0.98	0.07
2	**DenseNet-201**	0.99	1.00	0.99	0.99	0.005
3	**EfficientNet-B4**	0.93	0.96	0.93	0.93	0.37

Table 5.2: Performance Measures on the Corrupted Dataset

S No.	DCNNs	Noise	Accuracy	AUC	Precision	Recall	Loss
1	**Inception v3**	Gaussian	0.97	0.99	0.97	0.97	0.09
		Salt & Pepper	0.99	0.99	0.99	0.99	0.002
		Speckle	0.54	0.64	0.54	0.54	1.95
2	**DenseNet -201**	Gaussian	0.99	0.99	0.99	0.99	0.0007
		Salt & Pepper	0.99	0.99	0.99	0.99	0.004
		Speckle	0.83	0.91	0.83	0.83	0.39
3	**Efficient Net-B4**	Gaussian	0.77	0.85	0.77	0.77	1.02
		Salt & Pepper	1.00	1.00	1.00	1.00	0.0002

		Speckle	0.54	0.56	0.54	0.54	1.75

Table 5.3: Generalization Error

S No.	DCNN	A: Average Loss (Corrupted Dataset)	B: Average Loss (Validation Dataset)	Generalization Error(A-B)
1	**Inception v3**	0.68	0.07	0.61
2	**DenseNet-201**	0.13	0.005	0.12
3	**Efficient Net-B4**	0.92	0.37	0.55

5.5 Experimental Evaluation & Performance Analysis

This section highlights the visualizations of the observed results and also discusses the analysis of the proposed architecture.

5.5.1 Visualizations

Figures 5.7 and 5.8 depict the results of DCNNs on the validation and corrupted datasets. The generalization Error plot is visualized in Figure 5.9.

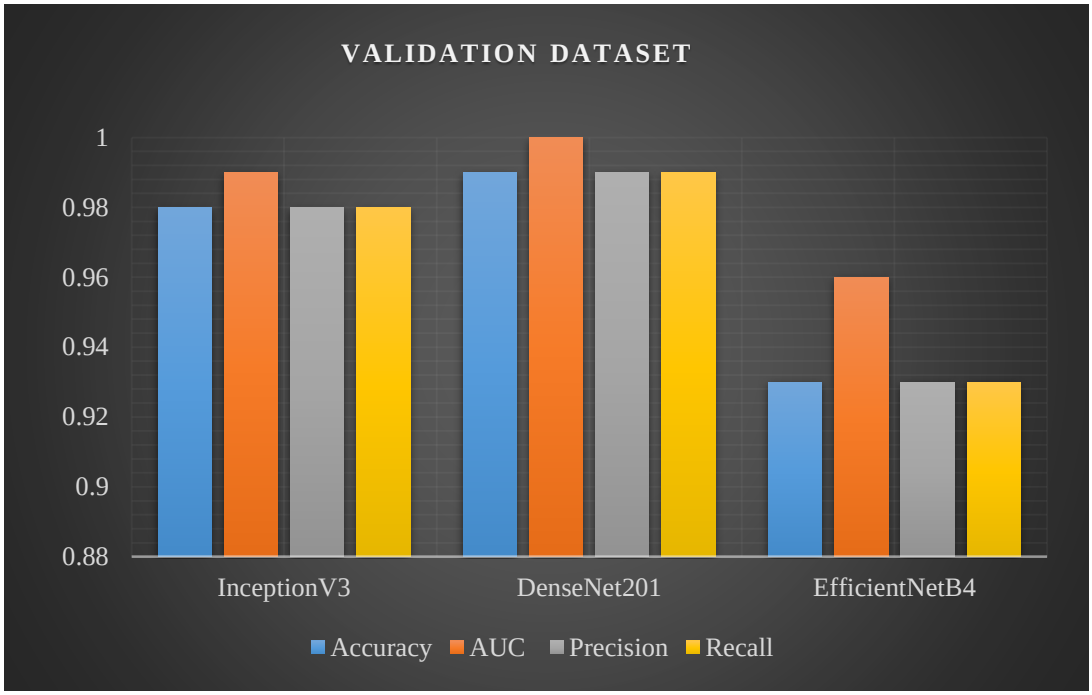

Fig. 5.7: Plots on Validation Data Set

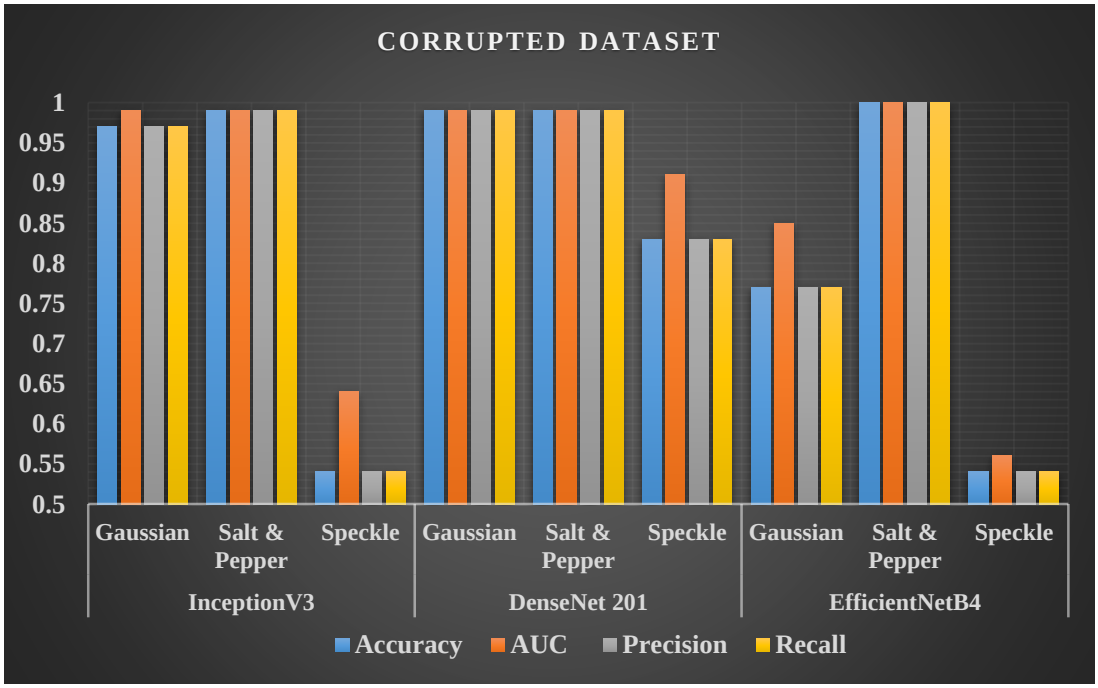

Fig. 5.8: Plots on Corrupted Dataset

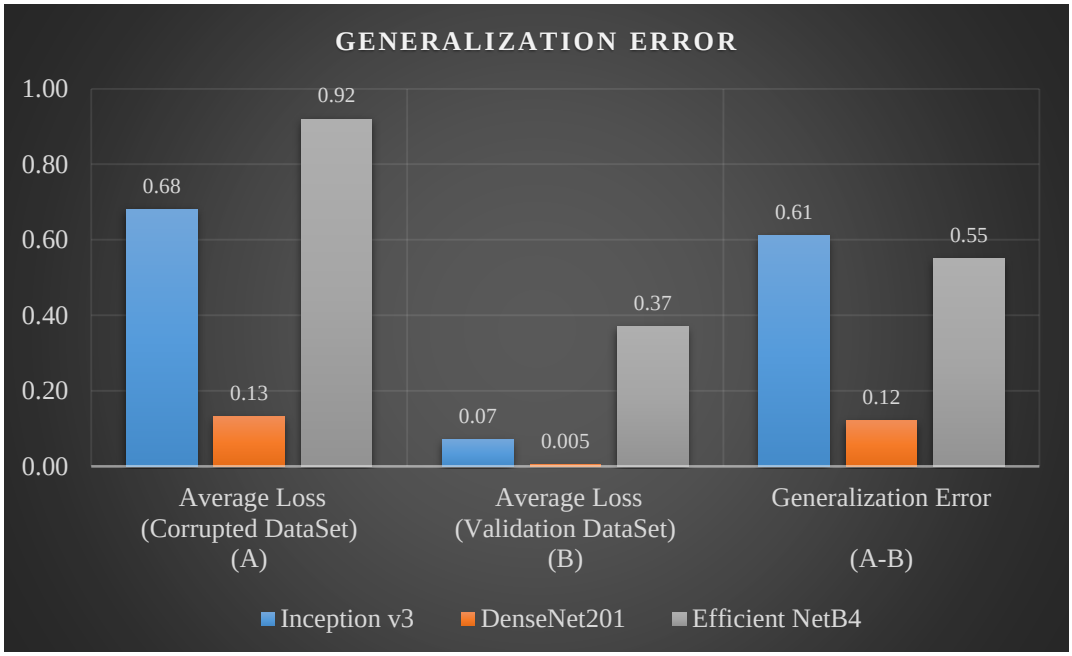

Fig. 5.9: Generalization Error

5.5.2 Analysis

This chapter proposed a framework for the estimation of Generation Error (GE) through noise induction on the CBIS-DDSM dataset. The Google Colaboratory functioned as the platform for the research findings. The dataset was spilt into training and validation in the proportion of 80:20. The model was trained in small batches of size 32 with Adam Optimizer. A learning rate of .0001 was adopted. Hidden layers of the model were trained with the ReLU activation whereas Softmax activation was applied in the final layer. Performance measures used for assessing the model's performance could be referred to in [13]. Tables 5.1, 5.2, and 5.3 show the outcomes of the proposed approach.

DenseNet-201 has shown superior performance with 99% accuracy on the Validation Dataset (Table 5.1). Validation Loss was also minimum for the DenseNet-201 model with a value of 0.005. On the other hand, Inception v3 and EfficientNet-B4 exhibit accuracies of 98% and 93% respectively with slightly high losses.

When the performance was evaluated on the Corrupted dataset, the outcomes are outlined in Table 5.2 and illustrated in Figure 5.2. The highest resistance to the Gaussian noise is shown by the DenseNet-201 model with a loss value of 0.0007.

EfficientNet-B4 has shown the highest tolerance to the Salt and pepper noise with a minimum loss of 0.0002. The distortion caused by Speckle noise is highest in Inceptionv3 and EfficientNet-B4 leading to lower accuracy and higher loss. Due to this deterioration, there is a significant decline in accuracy. Figures 5.7, 5.8, and 5.9 visualize performance measures graphically.

We calculated GE by estimating the difference in loss values on the Corrupted and Validation datasets (Table 5.3). DenseNet-201 exhibited a minimum GE of 0.12, on the other hand, Inception v3 and Efficient Net B4 have shown very high GE i.e. 0.61 and 0.55 respectively (Figure 5.5). From Table 5.2, we concluded that the generalizability shown by each DCNN on noisy data is different and it varies based on the category of noise being added to the dataset. The highest generalizability on the noisy data was depicted by DenseNet-201 with the least generalization error.

5.6 Chapter Summary

This chapter presents a design for assessing the generalizability of DCNNs. Gaussian, Salt and pepper, and Speckle noise were added to the CBIS-DDSM dataset. Generalizability was evaluated for three DCNNs - Inception Net v3, DenseNet-201, and EfficientNet-B4. The results were evaluated on the corrupted and validation data sets and then compared to evaluate the generalization error. Various performance measures were plotted to visualize the model's efficiency. The maximum distortion in the mammograms were caused by Speckle Noise with very high loss rates leading to low generalizability. Thus, it could be concluded that the proposed framework with DenseNet-201 has minimum generalization error and thus exhibits high generalizability on the noisy data.

PUBLICATION

The research addressed in this Chapter has been published in the following IEEE Conference:

- G. Chugh, S. Kumar and N. Singh, "A Framework for Generalization Error Evaluation in Deep Convolutional Neural Networks," *2023 IEEE Engineering Informatics*, Melbourne, Australia, 2023, pp. 1-7, doi: 10.1109/IEEECONF58110.2023.10520375.

CHAPTER 6

CONCLUSION, FUTURE PROSPECTIVE & SOCIAL IMPACT

6.1 Conclusion

Breast carcinoma is the premier category of deadliest cancer reported in females. There arises an urgency for early diagnosis and prognosis to lower the mortality rate. Medical image analysis using CAD has cropped up as an essential field for early diagnosis and prognosis in the healthcare domain. Although malignancy can't be proven without biopsy, early carcinoma detection using imaging modalities is an hour of need. Mammography continues to be used as the "gold standard" for breast carcinoma diagnosis owing to its widespread availability compared to others. Current research challenges suggest that technical and practical investigation is desperately needed to boost healthcare over the long term.

The first framework proposed for diagnosing and classifying breast carcinoma is named Transnet. In this dual model approach, two experiments were performed on the CBIS-DDSM dataset. The following deep neural networks were utilized- VGG-16, VGG-19, Mobile Net, ResNet-50, ResNet-152, and DenseNet-169. In the first experiment, namely Deep feature fusion with ML Classifier, pre-trained networks were deployed as feature extractors, and then the obtained attributes were provided to ML classifiers for classification. The second experiment, called Deep feature fusion with Neural Net classifiers, fine-tuned these networks for feature extraction and categorization. KNN and XGB classifiers perform best in the first approach yielding an accuracy of 100% on all the networks in the training phase. Conversely, ResNet-152 outperforms the other pre-trained networks by producing a 6% increase in accuracy on the test dataset. In the second experiment, Mobile Net, ResNet-50, and DenseNet-169 performed best in the training and testing phase with Accuracy and AUC of 97%(+-2%) and 0.97(+-0.02). The minimum loss, however, was exhibited by Mobile Net in both the train and test phases. The second approach performed better than the first, thus, improving all the

evaluation metrics. The results revealed that the proposed architecture performed remarkably well than the other revolutionary techniques.

Another framework proposed to enhance performance and utilize smaller datasets in the health domain is founded on the Multi Stage Transfer Learning Approach(MSTLA). Three mammography datasets were utilized: MIAS, In-Breast, and DDSM. The model is fine-tuned in three stages on separate datasets, and the optimized DCNN obtained in each stage is carried forward at the next stage. Two DCNNs are deployed for training the model – DenseNet-169 and ResNet-152. The results have shown that Stage 3 performs best compared to the other two stages, with DenseNet-169 having accuracy and AUC values of 100 and 1.0. The proposed approach could be employed for early-stage breast carcinoma diagnosis.

Another model proposed in this research is on the Generalizability of DNNs. Generalization is an approach to analyze how the model behaves on unseen data. Generalization Error(GE) measures the difference between training and testing errors. Gaussian, Salt and pepper, and Speckle noise were added to the CBIS-DDSM dataset. Generalizability was evaluated for three DCNNs - Inception Net v3, DenseNet-201, and EfficientNet-B4. DenseNet-201 performed remarkably well on Gaussian and Salt and pepper noise with a minimum loss rate i.e. .0007 for Gaussian noise. EfficientNet-B4 and Inception v3 have given the best results for Salt and Pepper noise with a minimum loss rate of .0002 and 0.002 respectively. The largest distortion in the mammograms was caused by Speckle Noise in Inception v3 and EfficientNet-B4 with very high loss rates leading to low generalizability. Generalization Error for Inception v3, DenseNet-201, and EfficientNet-B4 were 0.61, 0.12, and 0.55 respectively. Results have shown that the proposed framework with DenseNet-201 has minimum generalization error and thus exhibits high generalizability on the unseen i.e. noisy data.

Thus, the research presents efficient and optimal approaches for early-stage breast cancer diagnosis. The proposed strategies could be deployed in laboratories to assist doctors and pathologists in timely and precise diagnosis. Future perspectives of the proposed methodology include its implementation on several imaging techniques such as Ultrasound, MRI, CT, etc. The researchers can also evaluate the generalizability through various adversarial attacks on deep neural networks.

6.2 Applications & Future Prospects of the Proposed Research

Breast malignancy is climbing at a frightening pace globally. In contrast to developed nations, fatality counts are comparably high in low-wage and middle-wage nations[23]. Early diagnosis using CAD has become crucial in improving long-term survivability [27]. The symptomatic potentialities of training methods are impending the degree of personal competence using deep learning. Thus, the CAD paradigm of the "second opinion" tool is now being shifted to a more collaborative utility[29].

In the past few years, artificial intelligence has come up with advanced methods for the examination of medical pictures to aid radiologists at varying diagnosis stages[17]. DL is beneficial for handling intricate, heterogeneous, unorganized, and poorly annotated data [155]. Since DL eliminates the requirement for feature engineering, particularly when processing redundant data, the researchers extensively employ it in the majority of their investigations. Second, it's simple to adapt or modify current deep learning systems for use in new applications[26]. However, its real-time execution and deployment to clinics and hospitals is limited due to its computational and storage costs. The following are some crucial issues that require consideration when adopting deep learning for breast carcinoma diagnosis:

- **Using distinct imaging techniques and multiple modalities for breast malignancy**

In this study, we have used mammography, and the prospects include implementing the frameworks on other modalities such as Ultrasound, MRI, DBT, etc. As a result, to improve the models' capacity for classification and boost efficiency and reliability, the same patient might be diagnosed using various modalities.

- **Unsupervised approaches for breast tumor diagnosis**

The proposed research implemented supervised deep learning, i.e., CNN, that requires a large annotated dataset for breast carcinoma diagnosis. But in the health domain, the collection of labelled datasets is challenging, and most of the available datasets are unlabelled. Consequently, more research is needed to produce CAD systems that employ these unlabelled pictures, which are an essential data source.

• Less availability of dependable, comprehensive data sets from experienced physicians for breast malignancy

DL models need considerable training data. The unavailability of large datasets from experienced doctors is another issue that requires consideration. The generalizability of deep neural networks could be improved if they are tested on these datasets. Thus, the health industry could be improved if substantial annotated datasets from experienced physicians are made available to researchers[11].

• Integrating non-imaging information with imaging data

Till now, there are very few CAD systems that combine radiomic attributes with imaging data. Geras et al.[25] discussed that it is necessary to produce models that can incorporate image data with non-imaging characteristics, to diagnose malignancy in the initial stages.

• Imbalanced dataset

Another challenge for deep neural networks is class imbalance. Many available datasets are still imbalanced. A bias towards the more prevalent class could exist when models are trained on an imbalanced dataset and have a pessimistic effect on the classifier's behaviour [86]. Thus, the unavailability of balanced datasets is another issue that needs to be addressed.

• Generalizability and Generalization error of deeply convolutional neural networks

The word "generalization error" concerning DCNN is vague and unclear. Generally, we define generalizability as the network's behaviour on unfamiliar data. This study proposes a strategy for evaluating generalizability and generalization error by corrupting the trained dataset through different noise. We have utilized the mammography dataset. In the future, the researchers can also extend the proposed generalization error approach to other modalities. The generalization could also be assessed through different parameters such as various adversarial attacks. The authors could propose defense mechanisms for these attacks on deep neural networks.

• Less availability of models pre-trained on medical images

The existing pre-trained networks are trained on non-medical data. In recent years, very few researchers have focused on training a network on medical images and, thus, utilizing these networks further for transfer learning. If the dataset is too narrow, i.e., less than 1000 images, pre-trained networks experience overfitting issues. Thus, there arises a demand for domain-specific networks that are specially trained on large amounts of clinical pictures that when utilized would require less time, fewer resources, and are cost-effective.

6.3 Social Impact

Breast carcinoma is the deadliest malignancy in women, having surpassed cervical and lung cancers. Mortality cases due to cancer are increasing at an alarming rate and are particularly affecting younger age groups. Therefore, timely diagnosis is one of the most critical concerns that must be addressed globally since it can significantly enhance overall survival rates. Due to the complexity of medical images, manual examination is quite difficult. Manual processing of medical images faces three main issues - firstly, lack of availability of multiple pathologists at one location. In addition, the process of manually analyzing images is arduous and unpleasant. Lastly, the diagnosis of breast carcinoma heavily depends on the pathologists' expertise and domain knowledge.

CAD-formatted medical image processing has become a beneficial gadget that helps physicians label clinical images, thus enabling early diagnosis and treatment [8–10]. CAD systems require the processing of medical images; thus, extensive computational algorithms must be developed to process those images. Artificial Intelligence(AI) has emerged as the most promising field for various types of research in the current industries. Deep Learning and Machine learning, the subfields of AI, are giving tremendous results in each & every sector. We also use these applications in our daily lives, like scrolling the search engines, taking to digital assistants, playing innovative games, and using social media apps. Etc. In recent years DL and ML are also been widely used in the medical sector. These advanced technologies are helping doctors in the treatment, reducing the diagnosis time and thus saving patients' lives.

This study offers a useful paradigm for diagnosing and categorizing breast carcinoma at the initial stages. The condition must be diagnosed as soon as possible to lower the death toll. The proposed frameworks could be utilized to assist doctors and radiologists in providing timely diagnosis. Due to the memorization capability of deep neural networks, generalizability must also be addressed, which signifies how the model behaves on unseen data. This research also proposes a strategy to evaluate generalization error and thus compute the generalizability in DNNs. Deep learning has driven major advances in the analysis of medical pictures, achieving remarkable outcomes in several tasks. There is still a substantial barrier in the form of restricted accessibility to training data, especially in the healthcare field where obtaining data can be expensive and governed by privacy laws.

Thus, accuracy and workflow efficiency must be considered before implementing CAD systems in clinical practice, and secondly, to avoid improper use, user education is crucial to grasp the features and constraints of CAD systems.

REFERENCES

[1] Breastcancer, "Breast Cancer Facts and Statistics," *Breastcancer.org*, vol. 501, no. c, p. 3082851, 2023.

[2] American Cancer Society, "American Cancer Society | Information and Resources about for Cancer: Breast, Colon, Lung, Prostate, Skin," *Phytochemicals*. 2023.

[3] "Breast cancer." https://www.who.int/news-room/fact-sheets/detail/breast-cancer (accessed Apr. 20, 2024).

[4] American Cancer Society, "Breast Cancer | Breast Cancer Information & Overview | American Cancer Society," *American Cancer Society. Cancer Facts and Figures Atlanta, Ga: American Cancer Society*. pp. 1–50, 2021, [Online]. Available: https://www.cancer.org/cancer/types/breast-cancer.html.

[5] M. W. Huang, C. W. Chen, W. C. Lin, S. W. Ke, and C. F. Tsai, "SVM and SVM ensembles in breast cancer prediction," *PLoS One*, vol. 12, no. 1, pp. 1–14, 2017, doi: 10.1371/journal.pone.0161501.

[6] U.S. National Institutes of Health (NIH) and U.S. National Cancer Institute (NCI), "Definition of breast cancer - NCI Dictionary of Cancer Terms - NCI," *NCI Dictionary of Cancer Terms*. 2024, [Online]. Available: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/breast-cancer.

[7] I. Mihaylov, Ma. Nisheva, and D. Vassilev, *Machine Learning Techniques for Survival Time Prediction in Breast Cancer*, vol. 1. Springer International Publishing, 2018.

[8] R. J. Kate and R. Nadig, "Stage-specific predictive models for breast cancer survivability," *Int. J. Med. Inform.*, vol. 97, pp. 304–311, 2017, doi: 10.1016/j.ijmedinf.2016.11.001.

[9] W. M. Team, "Global cancer burden growing, amidst mounting need for services," *Saudi Medical Journal*, vol. 45, no. 3. pp. 326–327, 2024, [Online]. Available: https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing--amidst-mounting-need-for-services.

[10] Society American Cancer, "Global cancer facts and figures," *American Cancer Society*. pp. 1–48, 2024.

[11] G. Murtaza *et al.*, "Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges," *Artif. Intell. Rev.*, vol. 53, no. 3, pp. 1655–1720, 2020, doi: 10.1007/s10462-019-09716-5.

[12] T. G. Debelee, F. Schwenker, A. Ibenthal, and D. Yohannes, “Survey of deep learning in breast cancer image analysis,” *Evol. Syst.*, vol. 11, no. 1, pp. 143–163, 2020, doi: 10.1007/s12530-019-09297-2.

[13] G. Chugh, S. Kumar, and N. Singh, “Survey on Machine Learning and Deep Learning Applications in Breast Cancer Diagnosis,” *Cognitive Computation*. 2021, doi: 10.1007/s12559-020-09813-6.

[14] N. Khuriwal and N. Mishra, “Breast Cancer Diagnosis Using Deep Learning Algorithm,” *Proc. - IEEE 2018 Int. Conf. Adv. Comput. Commun. Control Networking, ICACCCN 2018*, pp. 98–103, 2018, doi: 10.1109/ICACCCN.2018.8748777.

[15] D. Selvathi and A. A. Poornila, “Deep Learning Techniques for Breast Cancer Detection Using Medical Image Analysis,” *Biol. Ration. Comput. Tech. Image Process. Appl.*, vol. 25, no. November, pp. 85–109, 2018, doi: 10.1007/978-3-319-61316-1.

[16] Z. Lai and H. Deng, “Medical image classification based on deep features extracted by deep model and statistic feature fusion with multilayer perceptron,” *Comput. Intell. Neurosci.*, vol. 2018, 2018, doi: 10.1155/2018/2061516.

[17] H. P. Chan, R. K. Samala, L. M. Hadjiiski, and C. Zhou, *Deep Learning in Medical Image Analysis*, vol. 1213. 2020.

[18] A. Jalalian, S. B. T. Mashohor, H. R. Mahmud, M. I. B. Saripan, A. R. B. Ramli, and B. Karasfi, “Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: A review,” *Clin. Imaging*, vol. 37, no. 3, pp. 420–426, 2013, doi: 10.1016/j.clinimag.2012.09.024.

[19] S. Sharma and P. Chaudhary, “Machine learning and deep learning,” *Quantum Comput. Artif. Intell. Train. Mach. Deep Learn. Algorithms Quantum Comput.*, pp. 71–84, 2023, doi: 10.1515/9783110791402-004.

[20] S. Suganyadevi, V. Seethalakshmi, and K. Balasamy, “A review on deep learning in medical image analysis,” *Int. J. Multimed. Inf. Retr.*, vol. 11, no. 1, pp. 19–38, 2022, doi: 10.1007/s13735-021-00218-1.

[21] A. Duggento *et al.*, “A random initialization deep neural network for discriminating malignant breast cancer lesions,” *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, vol. 2019, pp. 912–915, 2019, doi: 10.1109/EMBC.2019.8856740.

[22] A. Kebaili, J. Lapuyade-Lahorgue, and S. Ruan, “Deep Learning Approaches for Data Augmentation in Medical Imaging: A Review,” *J. Imaging*, vol. 9, no. 4, 2023, doi: 10.3390/jimaging9040081.

[23] N. I. R. Yassin, S. Omran, E. M. F. El Houby, and H. Allam, "Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review," *Comput. Methods Programs Biomed.*, vol. 156, pp. 25–45, 2018, doi: 10.1016/j.cmpb.2017.12.012.

[24] S. Turgut, M. Dagtekin, and T. Ensari, "Microarray breast cancer data classification using machine learning methods," *2018 Electr. Electron. Comput. Sci. Biomed. Eng. Meet. EBBT 2018*, pp. 1–3, 2018, doi: 10.1109/EBBT.2018.8391468.

[25] K. J. Geras, R. M. Mann, and L. Moy, "Artificial intelligence for mammography and digital breast tomosynthesis: Current concepts and future perspectives," *Radiology*, vol. 293, no. 2, pp. 246–259, 2019, doi: 10.1148/radiol.2019182627.

[26] Z. Hu, J. Tang, Z. Wang, K. Zhang, L. Zhang, and Q. Sun, "Deep learning for image-based cancer detection and diagnosis − A survey," *Pattern Recognit.*, vol. 83, pp. 134–149, 2018, doi: 10.1016/j.patcog.2018.05.014.

[27] G. Carneiro, J. Nascimento, and A. P. Bradley, "Automated Analysis of Unregistered Multi-View Mammograms with Deep Learning," *IEEE Trans. Med. Imaging*, vol. 36, no. 11, pp. 2355–2365, 2017, doi: 10.1109/TMI.2017.2751523.

[28] N. Akhtar and A. Mian, "Threat of Adversarial Attacks on Deep Learning in Computer Vision: A Survey," *IEEE Access*, vol. 6, pp. 14410–14430, 2018, doi: 10.1109/ACCESS.2018.2807385.

[29] J. R. Burt *et al.*, "Deep learning beyond cats and dogs: Recent advances in diagnosing breast cancer with deep neural networks," *Br. J. Radiol.*, vol. 91, no. 1089, 2018, doi: 10.1259/bjr.20170545.

[30] M. Saha and C. Chakraborty, "Her2Net: A deep framework for semantic segmentation and classification of cell membranes and nuclei in breast cancer evaluation," *IEEE Trans. Image Process.*, vol. 27, no. 5, pp. 2189–2200, 2018, doi: 10.1109/TIP.2018.2795742.

[31] "Breast Cancer: Statistics." https://www.cancer.net/cancer-types/breast-cancer/statistics (accessed Jul. 12, 2023).

[32] PBCR, "Trends of Breast Cancer in India," 2020. https://www.breastcancerindia.net/statistics/trends.html (accessed May 26, 2020).

[33] B. Sahiner *et al.*, "Deep learning in medical imaging and radiation therapy," *Med. Phys.*, vol. 46, no. 1, pp. e1–e36, Jan. 2019, doi: 10.1002/MP.13264.

[34] Saily, "Convolutional Neural Network: An Overview," *Analytics Vidyhya*, 2022. https://www.analyticsvidhya.com/blog/2022/01/convolutional-neural-network-an-

overview/ (accessed Jun. 11, 2022).

[35] S. Kulshrestha, “What Is A Convolutional Neural Network?,” *Developing an Image Classifier Using TensorFlow*. 2019, doi: 10.1007/978-1-4842-5572-8_6.

[36] L. Alzubaidi *et al.*, *Review of deep learning: concepts, CNN architectures, challenges, applications, future directions*, vol. 8, no. 1. Springer International Publishing, 2021.

[37] “Deep view on Transfer learning with Iamge classification Pytorch | by purnasai gudikandula | Medium.” https://purnasaigudikandula.medium.com/deep-view-on-transfer-learning-with-iamge-classification-pytorch-5cf963939575 (accessed Jun. 11, 2022).

[38] D. A. Ragab, O. Attallah, M. Sharkas, J. Ren, and S. Marshall, “A framework for breast cancer classification using Multi-DCNNs,” *Comput. Biol. Med.*, vol. 131, no. January, p. 104245, 2021, doi: 10.1016/j.compbiomed.2021.104245.

[39] Vishal Yathish, “Loss Functions and Their Use In Neural Networks,” *Towards Data Science*. 2022, [Online]. Available: https://towardsdatascience.com/loss-functions-and-their-use-in-neural-networks-a470e703f1e9.

[40] W. Worlberg, W. N. Street, and O. Mangasarian, “UCI Machine Learning Repository: Breast Cancer Wisconsin (Diagnostic) Data Set,” *http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%2528Diagnostic%2529*. 2011, Accessed: Dec. 11, 2020. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+(diagnostic).

[41] University of Aveiro, “Breast Cancer Digital Repository,” 2018. https://bcdr.eu/information/about (accessed Dec. 11, 2020).

[42] J. Suckling *et al.*, “Mammographic Image Analysis Society (MIAS) database v1.21,” Aug. 2015, Accessed: Dec. 12, 2020. [Online]. Available: https://www.repository.cam.ac.uk/handle/1810/250394.

[43] M. Heath, K. Bowyer, D. Kopans, R. Moore, and P. Kegelmeyer, “THE DIGITAL DATABASE FOR SCREENING MAMMOGRAPHY.”

[44] L. Spanhol, F., Oliveira, L. S., Petitjean, C. and Heutte, “Breast Cancer Histopathological Database (BreakHis) – Laboratório Visão Robótica e Imagem,” 2016. https://web.inf.ufpr.br/vri/databases/breast-cancer-histopathological-database-breakhis/ (accessed Dec. 12, 2020).

[45] Surveillance Research Program (SRP), “Accessing the Data - SEER Datasets.” https://seer.cancer.gov/data/access.html (accessed Dec. 12, 2020).

[46] I. C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, and J. S. Cardoso,

"INbreast: Toward a Full-field Digital Mammographic Database.," *Acad. Radiol.*, vol. 19, no. 2, pp. 236–248, 2012, doi: 10.1016/j.acra.2011.09.014.

[47] J. E. E. Oliveira, M. O. Gueld, A. de A. Araújo, B. Ott, and T. M. Deserno, "Toward a standard reference database for computer-aided mammography," in *Medical Imaging 2008: Computer-Aided Diagnosis*, 2008, vol. 6915, p. 69151Y, doi: 10.1117/12.770325.

[48] A. Osareh and B. Shadgar, "Machine learning techniques to diagnose breast cancer," *2010 5th Int. Symp. Heal. Informatics Bioinformatics, HIBIT 2010*, pp. 114–120, 2010, doi: 10.1109/HIBIT.2010.5478895.

[49] M. Karabatak and M. C. Ince, "An expert system for detection of breast cancer based on association rules and neural network," *Expert Syst. Appl.*, vol. 36, no. 2 PART 2, pp. 3465–3469, 2009, doi: 10.1016/j.eswa.2008.02.064.

[50] K. Goyal, S. Prakriti, A. Preeti, and K. Mukesh, *Comparative Analysis of Machine Learning Algorithms for Breast Cancer Prognosis*, vol. 46. Springer Singapore, 2019.

[51] T. A. Shaikh and A. Rashid, *Applying Machine Learning Algorithms for Early Diagnosis and Prediction of Breast Cancer Risk*, vol. 46. Springer Singapore, 2019.

[52] M. Gupta and B. Gupta, "An Ensemble Model for Breast Cancer Prediction Using Sequential Least Squares Programming Method (SLSQP)," *2018 11th Int. Conf. Contemp. Comput. IC3 2018*, pp. 1–3, 2018, doi: 10.1109/IC3.2018.8530572.

[53] M. Kashif, K. R. Malik, S. Jabbar, and J. Chaudhry, *Application of machine learning and image processing for detection of breast cancer*. Elsevier Inc., 2019.

[54] F. Liu and Brown Mackenzie, *Breast Cancer Recognition by Support Vector Machine Combined with Daubechies Wavelet Transform and Principal Component Analysis*, vol. 2018. Springer International Publishing, 2018.

[55] S. Boughorbel, R. Al-Ali, and N. Elkum, "Model comparison for breast cancer prognosis based on clinical data," *PLoS One*, vol. 11, no. 1, pp. 1–15, 2016, doi: 10.1371/journal.pone.0146413.

[56] R. Kleinlein and D. Riaño, "Persistence of data-driven knowledge to predict breast cancer survival," *Int. J. Med. Inform.*, vol. 129, no. June, pp. 303–311, 2019, doi: 10.1016/j.ijmedinf.2019.06.018.

[57] A. Sakai *et al.*, "A method for the automated classification of benign and malignant masses on digital breast tomosynthesis images using machine learning and radiomic features," *Radiol. Phys. Technol.*, vol. 13, no. 1, pp. 27–36, 2020, doi: 10.1007/s12194-019-00543-5.

[58] B. Zheng, S. W. Yoon, and S. S. Lam, “Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms,” *Expert Syst. Appl.*, vol. 41, no. 4 PART 1, pp. 1476–1482, 2014, doi: 10.1016/j.eswa.2013.08.044.

[59] H. Asri, H. Mousannif, H. Al Moatassime, and T. Noel, “Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis,” *Procedia Comput. Sci.*, vol. 83, no. Fams, pp. 1064–1069, 2016, doi: 10.1016/j.procs.2016.04.224.

[60] S. Ram and S. Gupta, *Building Machine Learning Based Diseases Diagnosis System Considering Various Features of Datasets*, vol. 841. Springer Singapore, 2019.

[61] M. A. Al-antari, M. A. Al-masni, M. T. Choi, S. M. Han, and T. S. Kim, “A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification,” *Int. J. Med. Inform.*, vol. 117, no. June, pp. 44–54, 2018, doi: 10.1016/j.ijmedinf.2018.06.003.

[62] M. A. Al-masni *et al.*, “Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system,” *Comput. Methods Programs Biomed.*, vol. 157, pp. 85–94, 2018, doi: 10.1016/j.cmpb.2018.01.017.

[63] H. Jung *et al.*, “Detection of masses in mammograms using a one-stage object detector based on a deep convolutional neural network,” *PLoS One*, vol. 13, no. 9, pp. 1–16, 2018, doi: 10.1371/journal.pone.0203355.

[64] F. F. Ting, Y. J. Tan, and K. S. Sim, “Convolutional neural network improvement for breast cancer classification,” *Expert Syst. Appl.*, vol. 120, pp. 103–115, 2019, doi: 10.1016/j.eswa.2018.11.008.

[65] J. Arevalo, F. A. González, R. Ramos-Pollán, J. L. Oliveira, and M. A. Guevara Lopez, “Representation learning for mammography mass lesion classification with convolutional neural networks,” *Comput. Methods Programs Biomed.*, vol. 127, pp. 248–257, 2016, doi: 10.1016/j.cmpb.2015.12.014.

[66] M. M. Jadoon, Q. Zhang, I. U. Haq, S. Butt, and A. Jadoon, “Three-Class Mammogram Classification Based on Descriptive CNN Features,” *Biomed Res. Int.*, vol. 2017, 2017, doi: 10.1155/2017/3640901.

[67] H. Li, S. Zhuang, D. ao Li, J. Zhao, and Y. Ma, “Benign and malignant classification of mammogram images based on deep learning,” *Biomed. Signal Process. Control*, vol. 51, pp. 347–354, 2019, doi: 10.1016/j.bspc.2019.02.017.

[68] S. Li *et al.*, “Computer-aided assessment of breast density: Comparison of supervised

deep learning and feature based statistical learning," *Phys Med Biol.*, 2018, doi: 10.1088/1361-6560/aa9f87.

[69] B. Mughal, N. Muhammad, and M. Sharif, "Adaptive hysteresis thresholding segmentation technique for localizing the breast masses in the curve stitching domain," *Int. J. Med. Inform.*, vol. 126, no. February, pp. 26–34, 2019, doi: 10.1016/j.ijmedinf.2019.02.001.

[70] F. Gao *et al.*, "SD-CNN: A shallow-deep CNN for improved breast cancer diagnosis," *Comput. Med. Imaging Graph.*, vol. 70, pp. 53–62, 2018, doi: 10.1016/j.compmedimag.2018.09.004.

[71] H. Cai *et al.*, "Breast Microcalcification Diagnosis Using Deep Convolutional Neural Network from Digital Mammograms," *Comput. Math. Methods Med.*, vol. 2019, 2019, doi: 10.1155/2019/2717454.

[72] D. Arefan, A. A. Mohamed, W. A. Berg, M. L. Zuley, J. H. Sumkin, and S. Wu, "Deep learning modeling using normal mammograms for predicting breast cancer risk," *Med. Phys.*, vol. 47, no. 1, pp. 110–118, 2020, doi: 10.1002/mp.13886.

[73] P. J. Sudharshan, C. Petitjean, F. Spanhol, L. E. Oliveira, L. Heutte, and P. Honeine, "Multiple instance learning for histopathological breast cancer image classification," *Expert Syst. Appl.*, vol. 117, pp. 103–111, 2019, doi: 10.1016/j.eswa.2018.09.049.

[74] R. K. Samala, H. P. Chan, L. M. Hadjiiski, M. A. Helvie, K. Cha, and C. Richter, "Multi-task transfer learning deep convolutional neural network: application to computer-aided diagnosis of breast cancer on mammograms," *Phys Med Biol.*, p. 8894-8908, 2017, doi: 10.1088/1361-6560/aa93d4.

[75] M. Yousefi, A. Krzyżak, and C. Y. Suen, "Mass detection in digital breast tomosynthesis data using convolutional neural networks and multiple instance learning," *Comput. Biol. Med.*, vol. 96, no. February, pp. 283–293, 2018, doi: 10.1016/j.compbiomed.2018.04.004.

[76] M. H. Yap *et al.*, "Automated Breast Ultrasound Lesions Detection Using Convolutional Neural Networks," *IEEE J. Biomed. Heal. Informatics*, vol. 22, no. 4, pp. 1218–1226, 2018, doi: 10.1109/JBHI.2017.2731873.

[77] X. Qi *et al.*, "Automated diagnosis of breast ultrasonography images using deep neural networks," *Med. Image Anal.*, vol. 52, pp. 185–198, 2019, doi: 10.1016/j.media.2018.12.006.

[78] T. Araujo *et al.*, "Classification of Breast Cancer Based on Histology Images Using Convolutional Neural Networks," *PLoS One*, vol. 6, pp. 24680–24693, 2017, doi:

10.1371/journal.pone.0177544.

[79] K. Roy, D. Banik, D. Bhattacharjee, and M. Nasipuri, “Patch-based system for Classification of Breast Histology images using deep learning,” *Comput. Med. Imaging Graph.*, vol. 71, pp. 90–103, 2019, doi: 10.1016/j.compmedimag.2018.11.003.

[80] Y. Li, J. Wu, and Q. Wu, “Classification of Breast Cancer Histology Images Using Multi-Size and Discriminative Patches Based on Deep Learning,” *IEEE Access*, vol. 7, pp. 21400–21408, 2019, doi: 10.1109/ACCESS.2019.2898044.

[81] E. Deniz, A. Şengür, Z. Kadiroğlu, Y. Guo, V. Bajaj, and Ü. Budak, “Transfer learning based histopathologic image classification for breast cancer detection,” *Heal. Inf. Sci. Syst.*, vol. 6, no. 1, 2018, doi: 10.1007/s13755-018-0057-x.

[82] N. Brancati, G. De Pietro, M. Frucci, and D. Riccio, “A Deep Learning Approach for Breast Invasive Ductal Carcinoma Detection and Lymphoma Multi-Classification in Histological Images,” *IEEE Access*, vol. 7, no. c, pp. 44709–44720, 2019, doi: 10.1109/ACCESS.2019.2908724.

[83] Y. Jiang, L. Chen, H. Zhang, and X. Xiao, “Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module,” *PLoS One*, vol. 14, no. 3, pp. 1–21, 2019, doi: 10.1371/journal.pone.0214587.

[84] A. Al Nahid, M. A. Mehrabi, and Y. Kong, “Histopathological breast cancer image classification by deep neural network techniques guided by local clustering,” *Biomed Res. Int.*, vol. 2018, 2018, doi: 10.1155/2018/2362108.

[85] R. Yan *et al.*, *Integration of Multimodal Data for Breast Cancer Classification Using a Hybrid Deep Learning Method*, vol. 11643 LNCS. Springer International Publishing, 2019.

[86] S. Sharma and R. Mehra, “Conventional Machine Learning and Deep Learning Approach for Multi-Classification of Breast Cancer Histopathology Images—a Comparative Insight,” *J. Digit. Imaging*, 2020, doi: 10.1007/s10278-019-00307-y.

[87] V. Bevilacqua, A. Brunetti, A. Guerriero, G. F. Trotta, M. Telegrafo, and M. Moschetta, “A performance comparison between shallow and deeper neural networks supervised classification of tomosynthesis breast lesions images,” *Cogn. Syst. Res.*, vol. 53, pp. 3–19, 2019, doi: 10.1016/j.cogsys.2018.04.011.

[88] X. Zhang *et al.*, “Classification of whole mammogram and tomosynthesis images using deep convolutional neural networks,” *IEEE Trans. Nanobioscience*, vol. 17, no. 3, pp. 237–242, 2018, doi: 10.1109/TNB.2018.2845103.

[89] R. K. Samala, H. P. Chan, L. Hadjiiski, M. A. Helvie, J. Wei, and K. Cha, “Mass

detection in digital breast tomosynthesis: Deep convolutional neural network with transfer learning from mammography," *Med. Phys.*, vol. 43, no. 12, pp. 6654–6666, 2016, doi: 10.1118/1.4967345.

[90] M. Fan, Y. Li, S. Zheng, W. Peng, W. Tang, and L. Li, "Computer-aided detection of mass in digital breast tomosynthesis using a faster region-based convolutional neural network," *Methods*, vol. 166, pp. 103–111, 2019, doi: 10.1016/j.ymeth.2019.02.010.

[91] D. Zhang, L. Zou, X. Zhou, and F. He, "Integrating Feature Selection and Feature Extraction Methods with Deep Learning to Predict Clinical Outcome of Breast Cancer," *IEEE Access*, vol. 6, no. 8, pp. 28936–28944, 2018, doi: 10.1109/ACCESS.2018.2837654.

[92] X. Zhang *et al.*, "Extracting comprehensive clinical information for breast cancer using deep learning methods," *Int. J. Med. Inform.*, vol. 132, p. 103985, 2019, doi: 10.1016/j.ijmedinf.2019.103985.

[93] P. Agarwal, A. Yadav, and P. Mathur, "Breast Cancer Prediction on BreakHis Dataset Using Deep CNN and Transfer Learning Model," *Lect. Notes Networks Syst.*, vol. 238, pp. 77–88, 2022, doi: 10.1007/978-981-16-2641-8_8.

[94] N. M. Hassan, S. Hamad, and K. Mahar, *Mammogram breast cancer CAD systems for mass detection and classification: a review*, vol. 81, no. 14. Multimedia Tools and Applications, 2022.

[95] G. Ayana, J. Park, J. W. Jeong, and S. W. Choe, "A Novel Multistage Transfer Learning for Ultrasound Breast Cancer Image Classification," *Diagnostics*, vol. 12, no. 1. 2022, doi: 10.3390/diagnostics12010135.

[96] M. Rana and M. Bhushan, "Machine learning and deep learning approach for medical image analysis: diagnosis to detection," *Multimed. Tools Appl.*, pp. 26731–26769, 2022, doi: 10.1007/s11042-022-14305-w.

[97] K. Vijayakumar, V. J. Kadam, and S. K. Sharma, "Breast cancer diagnosis using multiple activation deep neural network," *Concurr. Eng. Res. Appl.*, vol. 29, no. 3, pp. 275–284, Sep. 2021, doi: 10.1177/1063293X211025105.

[98] E. M. Senan, F. W. Alsaade, M. I. A. Al-Mashhadani, T. H. H. Aldhyani, and M. H. Al-Adhaileh, "Classification of histopathological images for early detection of breast cancer using deep learning," *J. Appl. Sci. Eng.*, vol. 24, no. 3, pp. 323–329, 2021, doi: 10.6180/JASE.202106_24(3).0007.

[99] W. Sun, T. L. B. Tseng, J. Zhang, and W. Qian, "Enhancing deep convolutional neural network scheme for breast cancer diagnosis with unlabeled data," *Comput. Med.*

Imaging Graph., vol. 57, pp. 4–9, 2017, doi: 10.1016/j.compmedimag.2016.07.004.

[100] D. M. Vo, N. Q. Nguyen, and S. W. Lee, "Classification of breast cancer histology images using incremental boosting convolution networks," *Inf. Sci. (Ny).*, vol. 482, pp. 123–138, 2019, doi: 10.1016/j.ins.2018.12.089.

[101] L. Li *et al.*, "Multi-task Deep Learning for Fine-Grained Classification and grading in Breast Cancer Histopathological Images," *Multimed. Tools Appl.*, 2018, doi: 10.1007/s11042-018-6970-9.

[102] R. K. Samala, H. P. Chan, L. Hadjiiski, M. A. Helvie, C. D. Richter, and K. H. Cha, "Breast cancer diagnosis in digital breast tomosynthesis: Effects of training sample size on multi-stage transfer learning using deep neural nets," *IEEE Trans. Med. Imaging*, vol. 38, no. 3, pp. 686–696, 2019, doi: 10.1109/TMI.2018.2870343.

[103] N. Antropova, B. Q. Huynh, and M. L. Giger, "A Deep Feature Fusion Methodology for Breast Cancer Diagnosis Demonstrated on Three Imaging Modality Datasets," *Med Phys.*, vol. 44, p. 5162-5171, 2017, doi: 10.1002/mp.12453.

[104] G. Chugh, S. Kumar, and N. Singh, "A Framework for Generalization Error Evaluation in Deep Convolutional Neural Networks," *2023 IEEE Eng. Informatics*, pp. 1–7, 2023, doi: 10.1109/IEEECONF58110.2023.10520375.

[105] R. K. Samala, H. P. Chan, L. M. Hadjiiski, M. A. Helvie, and C. D. Richter, "Generalization error analysis for deep convolutional neural network with transfer learning in breast cancer diagnosis," *Phys. Med. Biol.*, vol. in press, 2020, doi: 10.1088/1361-6560/ab82e8.

[106] L. Garrucho, K. Kushibar, S. Jouide, O. Diaz, L. Igual, and K. Lekadir, "Domain generalization in deep learning based mass detection in mammography: A large-scale multi-center study," *Artif. Intell. Med.*, vol. 132, 2022, doi: 10.1016/j.artmed.2022.102386.

[107] O. Joshua, J. S. Owotogbe, T. S. Ibiyemi, and B. A. Adu, "A Comprehensive Review On Various Types of Noise in Image Processing The Role of Robust ICT in Fostering Agricultural Extension, Rural Development and Food Security View project Image Processing View project A Comprehensive Review On Various Types of Nois," *Artic. Int. J. Sci. Eng. Res.*, no. December, 2019, [Online]. Available: http://www.ijser.org.

[108] D. Jakubovitz, R. Giryes, and M. R. D. Rodrigues, "Generalization Error in Deep Learning," *Appl. Numer. Harmon. Anal.*, pp. 153–193, 2019, doi: 10.1007/978-3-319-73074-5_5.

[109] H. Rajaguru and S. R. Sannasi Chakravarthy, "Effiient denoising framework for

mammogram images with a new impulse detector and non-local means," *Asian Pacific J. Cancer Prev.*, vol. 21, no. 1, pp. 179–183, 2020, doi: 10.31557/APJCP.2020.21.1.179.

[110] X. Ma *et al.*, "Understanding adversarial attacks on deep learning based medical image analysis systems," *Pattern Recognit.*, vol. 110, 2021, doi: 10.1016/j.patcog.2020.107332.

[111] Z. Li *et al.*, "Domain Generalization for Mammography Detection via Multi-style and Multi-view Contrastive Learning," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12907 LNCS, no. Dl, pp. 98–108, 2021, doi: 10.1007/978-3-030-87234-2_10.

[112] R. S. Lee, F. Gimenez, A. Hoogi, K. K. Miyake, M. Gorovoy, and D. L. Rubin, "Data Descriptor: A curated mammography data set for use in computer-aided detection and diagnosis research," *Sci. Data*, vol. 4, Dec. 2017, doi: 10.1038/SDATA.2017.177.

[113] G. Chugh, S. Kumar, and N. Singh, "TransNet: a comparative study on breast carcinoma diagnosis with classical machine learning and transfer learning paradigm," *Multimed. Tools Appl.*, no. 0123456789, 2023, doi: 10.1007/s11042-023-16938-x.

[114] K. Simonyan and A. Zisserman, "VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION," 2015. [Online]. Available: http://www.robots.ox.ac.uk/.

[115] V. Paras, "VGGNet-16 Architecture: A Complete Guide," *Kaggle*. 2020, [Online]. Available: https://www.kaggle.com/code/blurredmachine/vggnet-16-architecture-a-complete-guide.

[116] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition." [Online]. Available: http://image-net.org/challenges/LSVRC/2015/.

[117] N. A. Al-Humaidan and M. Prince, "A Classification of Arab Ethnicity Based on Face Image Using Deep Learning Approach," *IEEE Access*, vol. 9, no. March, pp. 50755–50766, 2021, doi: 10.1109/ACCESS.2021.3069022.

[118] P. T. Hr and S. Ložnjak1, "Automobile Classification Using Transfer Learning on Resnet Neural Network Architecture Klasifikacija Automobila Koristeći Transferirano Učenje Na Resnet Arhitekturi Neuronske Mreže," *Polytech. Des.*, vol. 8, no. 1, 2020, doi: 10.19279/TVZ.PD.2020-8-1-18.

[119] "MobileNet Convolutional neural network Machine Learning Algorithms | Analytics Vidhya." https://medium.com/analytics-vidhya/image-classification-with-mobilenet-cc6fbb2cd470 (accessed Jun. 11, 2022).

[120] S. Phiphiphatphaisit and O. Surinta, “Food Image Classification with Improved MobileNet Architecture and Data Augmentation,” *ACM Int. Conf. Proceeding Ser.*, no. April, pp. 51–56, 2020, doi: 10.1145/3388176.3388179.

[121] M. Steen, S. Downe, N. Bamford, and L. Edozien, “DenseNet:Densely Connected Convolutional Networks arXiv:1608.06993v5,” *Arxiv*, vol. 28, no. 4, pp. 362–371, 2018.

[122] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, and M. Kaur, “Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning,” *J. Biomol. Struct. Dyn.*, vol. 39, no. 15, pp. 5682–5689, 2021, doi: 10.1080/07391102.2020.1788642.

[123] S. Uddin, I. Haque, H. Lu, M. A. Moni, and E. Gide, “Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction,” *Sci. Rep.*, vol. 12, no. 1, pp. 1–11, 2022, doi: 10.1038/s41598-022-10358-x.

[124] S. Huang, C. A. I. Nianguang, P. Penzuti Pacheco, S. Narandes, Y. Wang, and X. U. Wayne, “Applications of support vector machine (SVM) learning in cancer genomics,” *Cancer Genomics and Proteomics*, vol. 15, no. 1, pp. 41–51, 2018, doi: 10.21873/cgp.20063.

[125] K. Fawagreh, M. M. Gaber, and E. Elyan, “Random forests: From early developments to recent advancements,” *Syst. Sci. Control Eng.*, vol. 2, no. 1, pp. 602–609, 2014, doi: 10.1080/21642583.2014.956265.

[126] A. Kumar, “AdaBoost Algorithm: Boosting Algorithm in Machine Learning,” *Great Learning*. 2022.

[127] R. E. Schapire, “The Boosting Approach to Machine Learning: An Overview,” pp. 149–171, 2003, doi: 10.1007/978-0-387-21579-2_9.

[128] C. Bentéjac, A. Csörgő, and G. Martínez-Muñoz, *A comparative analysis of gradient boosting algorithms*, vol. 54, no. 3. Springer Netherlands, 2021.

[129] “XGBoost.” [Online]. Available: https://www.nvidia.com/en-us/glossary/xgboost/.

[130] H. Shu *et al.*, “A deep learning approach to re-create raw full-field digital mammograms for breast density and texture analysis,” *Radiol. Artif. Intell.*, vol. 3, no. 4, 2021, doi: 10.1148/ryai.2021200097.

[131] “Segmenting Abnormalities in Mammograms (Part 2 of 3) | by Cleon W | Towards Data Science.” https://towardsdatascience.com/can-you-find-the-breast-tumours-part-2-of-3-1d43840707fc (accessed Jun. 11, 2022).

[132] F. Alrasheedi, X. Zhong, and P. C. Huang, “Padding Module: Learning the Padding in Deep Neural Networks,” *IEEE Access*, vol. 11, no. January, pp. 7348–7357, 2023, doi: 10.1109/ACCESS.2023.3238315.

[133] Y. J. Tan, K. S. Sim, and F. F. Ting, “Breast cancer detection using convolutional neural networks for mammogram imaging system,” *Proceeding 2017 Int. Conf. Robot. Autom. Sci. ICORAS 2017*, vol. 2018-March, pp. 1–5, Mar. 2018, doi: 10.1109/ICORAS.2017.8308076.

[134] E. Wu, K. Wu, D. Cox, and W. Lotter, “Conditional infilling GANs for data augmentation in mammogram classification,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11040 LNCS, pp. 98–106, 2018, doi: 10.1007/978-3-030-00946-5_11.

[135] D. A. Ragab, M. Sharkas, S. Marshall, and J. Ren, “Breast cancer detection using deep convolutional neural networks and support vector machines,” *PeerJ*, vol. 2019, no. 1, p. e6201, Jan. 2019, doi: 10.7717/PEERJ.6201/TABLE-8.

[136] H. N. KHAN, A. R. SHAHID, B. RAZA, A. H. DAR, and A. H. ALQUHAYZ, “Multi-View Feature Fusion Based Four Views Model for Mammogram Classification Using Convolutional Neural Network,” *IEEE Access*, 2019. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8897609 (accessed Jun. 14, 2022).

[137] R. Song, T. Li, and Y. Wang, “Mammographic Classification Based on XGBoost and DCNN with Multi Features,” *IEEE Access*, vol. 8, pp. 75011–75021, 2020, doi: 10.1109/ACCESS.2020.2986546.

[138] H. Zhang *et al.*, “DE-Ada * : A Novel Model for Breast Mass Classification Using Cross-modal Pathological Semantic Mining and Organic Integration of Multi-feature Fusions,” 2020, Accessed: Jun. 14, 2022. [Online]. Available: https://www.elsevier.com/open-access/userlicense/1.0/.

[139] T. Choudhary, V. Mishra, A. Goswami, and J. Sarangapani, “A transfer learning with structured filter pruning approach for improved breast cancer classification on point-of-care devices,” *Comput. Biol. Med.*, vol. 134, no. April, p. 104432, 2021, doi: 10.1016/j.compbiomed.2021.104432.

[140] M. Alruwaili and W. Gouda, “Automated Breast Cancer Detection Models Based on Transfer Learning,” *Sensors*, vol. 22, no. 3, 2022, doi: 10.3390/s22030876.

[141] N. Mobark, S. Hamad, and S. Z. Rida, “CoroNet: Deep Neural Network-Based End-to-End Training for Breast Cancer Diagnosis,” *Appl. Sci.*, vol. 12, no. 14, 2022, doi:

10.3390/app12147080.

[142] R. Sujatha, J. M. Chatterjee, A. Angelopoulou, E. Kapetanios, P. N. Srinivasu, and D. J. Hemanth, “A transfer learning-based system for grading breast invasive ductal carcinoma,” *IET Image Process.*, no. September 2022, pp. 1979–1990, 2022, doi: 10.1049/ipr2.12660.

[143] Y. Bai *et al.*, “Recognizing breast tumors based on mammograms combined with pre-trained neural networks,” *Multimed. Tools Appl.*, pp. 27989–28008, 2023, doi: 10.1007/s11042-023-14708-3.

[144] Y. Sahu, A. Tripathi, R. K. Gupta, P. Gautam, R. K. Pateriya, and A. Gupta, “A CNN-SVM based computer aided diagnosis of breast Cancer using histogram K-means segmentation technique,” *Multimed. Tools Appl.*, pp. 14055–14075, 2022, doi: 10.1007/s11042-022-13807-x.

[145] S. Khalil *et al.*, “Enhancing Ductal Carcinoma Classification Using Transfer Learning with 3D U-Net Models in Breast Cancer Imaging,” *Appl. Sci.*, vol. 13, no. 7, pp. 1–20, 2023, doi: 10.3390/app13074255.

[146] G. Chugh, S. Kumar, and N. Singh, “MSTLA: Multi-Stage Transfer Learning Approach for Breast Carcinoma Diagnosis,” *2023 Int. Conf. Adv. Comput. Comput. Technol. InCACCT 2023*, pp. 509–514, 2023, doi: 10.1109/InCACCT57535.2023.10141697.

[147] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the Inception Architecture for Computer Vision,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 2818–2826, 2016, doi: 10.1109/CVPR.2016.308.

[148] R. Tiwari, “How to use Inception Model for Image recognition.” 2021, [Online]. Available: https://www.indusmic.com/post/how-to-use-inception-model-for-image-recognition.

[149] M. Tan and Q. V. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” *36th Int. Conf. Mach. Learn. ICML 2019*, vol. 2019-June, pp. 10691–10700, 2019.

[150] B. Baheti, S. Innani, S. Gajre, and S. Talbar, “Eff-UNet : A Novel Architecture for Semantic Segmentation in Unstructured Environment Eff-UNet : A Novel Architecture for Semantic Segmentation in Unstructured Environment,” no. June, 2020, doi: 10.1109/CVPRW50498.2020.00187.

[151] C. Y. Zhu *et al.*, “A Deep Learning Based Framework for Diagnosing Multiple Skin

Diseases in a Clinical Environment," *Front. Med.*, vol. 8, no. April, pp. 1–13, 2021, doi: 10.3389/fmed.2021.626369.

[152] Joel Marina Z. *et al.*, "Adversarial Attack Vulnerability of Deep Learning Models for Oncologic Images," pp. 1–20, 2021.

[153] N. E. Khalifa, M. Loey, and S. Mirjalili, "A comprehensive survey of recent trends in deep learning for digital images augmentation," *Artif. Intell. Rev.*, vol. 55, no. 3, pp. 2351–2377, 2022, doi: 10.1007/s10462-021-10066-4.

[154] S. Pradeep and P. Nirmaladevi, "A Review on Speckle Noise Reduction Techniques in Ultrasound Medical images based on Spatial Domain, Transform Domain and CNN Methods," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1055, no. 1, p. 012116, 2021, doi: 10.1088/1757-899x/1055/1/012116.

[155] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, "Deep learning for healthcare: Review, opportunities and challenges," *Brief. Bioinform.*, vol. 19, no. 6, pp. 1236–1246, 2017, doi: 10.1093/bib/bbx044.

Cognitive Computation (2021) 13:1451–1470
https://doi.org/10.1007/s12559-020-09813-6

Survey on Machine Learning and Deep Learning Applications in Breast Cancer Diagnosis

Gunjan Chugh[1] · Shailender Kumar[1] · Nanhay Singh[2]

Received: 4 September 2020 / Accepted: 28 December 2020 / Published online: 30 January 2021

Abstract
Cancer is a fatal disease caused due to the undesirable spread of cells. Breast carcinoma is the most invasive tumors and is the main reason for cancer deaths in females. Therefore, early diagnosis and prognosis have become necessary to increase survivability and reduce death rates in the long run. New artificial intelligence technologies are assisting radiologists in medical image scrutiny, thereby improving cancer patients' status. This survey enrolls peer-reviewed, newly developed computer-aided diagnosis (CAD) systems implementing machine learning (ML) and deep learning (DL) techniques for diagnosing breast carcinoma, compares them with previously established methods, and provides technical details with the pros and cons for each model. We also discuss some open issues, research gaps, and future research directions for the advanced CAD models in medical image analysis. Over the past decade, machine learning and deep learning have emerged as a subfield of artificial intelligence (AI), whose healthcare industry applications have provided excellent results with reduced cost and improved efficiency. This survey analyzes different classifiers of machine learning and deep learning approaches for breast cancer diagnosis. Results from previous studies proved that deep learning outperforms conventional machine learning for diagnosing breast carcinoma when the dataset is broad. Research gaps from the recent studies depict that practical and scientific research is an urgent necessity for improving healthcare in the long run.

Keywords Breast cancer · Machine learning · Deep learning · Convolution neural network (CNN) · Computer-aided diagnosis (CAD)

Introduction

Our bodies have millions of cells. Cancer starts when cellular changes cause these cells to grow excessively, creating a lump called a primary tumor [1]. According to data reported in Cancer Statistics, 2018 [2], approximately 2.25 million people live with this disease. New registered cancer patients exceed 1,157,294, and nearly 784, 821 cancer-related deaths are reported annually. The risk of dying from cancer is approximately 7.34% in males and 6.28% in females. In males, 25% of cancer deaths were reported for oral cavity and lung cancer, whereas breast and oral cavity cancer accounts for 25% of cancers in females [2].

Figure 1 shows a detailed description of cancer statistics for the year 2018. As we can see, breast cancer is the most frequent carcinoma and constitutes 14% of all cancers in women. Globocan 2018 study [2] reported approximately 162,468 new breast cancer cases and 87,090 deaths.

Incidences of breast carcinoma are increasing at a startling rate in India, particularly in younger age groups—30s and 40s (Fig. 2). Women aged 20–59 years are most often affected by breast cancer [4]. In India, detection occurs very late, which is the main reason for low survival rates for breast carcinoma [5]. Over the last few decades, cancer-related research is being performed. Scientists have applied different methods to predict disease before they cause symptoms. Therefore, accurate diagnosis and prognosis of breast cancer are regarded as exciting and challenging tasks for physicians in the medical and healthcare communities [6].

✉ Shailender Kumar
shailenderkumar@dce.ac.in

[1] Department of Computer Science and Engg, Delhi Technological University, Delhi, India

[2] Department of Computer Science, Ambedkar Institute of Advanced Communication Technologies and Research, Delhi, India

Multimedia Tools and Applications
https://doi.org/10.1007/s11042-023-16938-x

TransNet: a comparative study on breast carcinoma diagnosis with classical machine learning and transfer learning paradigm

Gunjan Chugh[1] · Shailender Kumar[1] · Nanhay Singh[2]

Received: 20 February 2023 / Revised: 7 July 2023 / Accepted: 11 September 2023

Abstract
Breast Carcinoma is a deadly disease; therefore, timely diagnosis is one of the most critical concerns that must be addressed globally since it can significantly enhance overall survival rates. Currently, Medical Imaging relies on Machine Learning(ML) and Deep Learning(DL) for accurate and early identification of diseases. In this article, a framework is proposed for diagnosing & classifying breast tumors using deep learning approaches. We have performed two experiments on the CBIS-DDSM (Curated Breast Imaging Subset of Digital Database for Screening Mammography) dataset. In the first approach, i.e., Deep feature extraction with ML classifier head, Deep Convolutional Neural Network(DCNN) models such as VGG16, VGG19, Res Net 50, and Res Net 152 are deployed as feature extractors, and the obtained features are utilized for training conventional machine learning classifiers. The second approach, called Deep Learning feature extraction with a neural network classifier, exploits Mobile Net, VGG16, VGG19, ResNet50, Res Net 152, and, Dense Net 169 for feature extraction and categorization. The results show that in the first case, Random Forest (RF) and XG Boost (XGB) Classifier perform best with 100% accuracy on the training set, whereas Support Vector Machine (SVM) and XGB exhibit 95%(+-5%) on the Test dataset for all the models. In the second approach, Mobile Net, ResNet50, and Dense Net 169 outperform the other models with an accuracy of 97%(+-2%) for both the Training and Test sets. The evaluated results have shown that the second approach depicts an increase in accuracy by 4%.

Keywords Breast Carcinoma · Computer-Aided diagnosis · Deep Convolutional Neural Network · Deep-learning

✉ Gunjan Chugh
chugh.gunjan8917@gmail.com

[1] Department of Computer Science and Engineering, Delhi Technological University, Delhi, India

[2] Department of Computer Science and Engineering, Netaji Subhas University of Technology (East Campus), Delhi, India

Published online: 25 September 2023

2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT) | 979-8-3503-9648-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/InCACCT57535.2023.10141697

MSTLA: Multi-Stage Transfer Learning Approach for Breast Carcinoma Diagnosis

Gunjan Chugh
Dept. of Computer Science & Engg.
Delhi Technological University
New Delhi, India
chugh.gunjan8917@gmail.com

Shailender Kumar
Dept. of Computer Science & Engg.
Delhi Technological University
New Delhi, India
shailenderkumar@dce.ac.in

Nanhay Singh
Dept. of Computer Science & Engg.
Netaji Subhas University of Technology (East Campus)
New Delhi, India
nsingh1973@gmail.com

***Abstract*- In the recent era, Breast Carcinoma has been observed as a fatal disease in women. In India and other developing countries, survival rates are very low because detection occurs very late. Computer-aided diagnosis(CAD) has emerged as a tool for helping doctors and experts with early detection and diagnosis. Medical Image Processing relies upon processing medical images to segment and categorize several diseases. Machine Learning(ML) and Deep Learning(DL) have shown tremendous success in recent years in various real-life applications, including speech recognition, face recognition, autonomous vehicles, etc. Transfer Learning is a technique where features learned from one domain are transferred to another. This approach is generally followed when the data required for training a model is unavailable in abundance. In this work, we have designed a model for the early diagnosis and categorization of breast malignancy using the Multi-Stage Transfer Learning Approach(MSTLA). DenseNet169 and ResNet152 are utilized for three stage transfer learning strategy. The results show that both DenseNet169 and ResNet152 performed remarkably with an accuracy of 100% and 99% in the third stage of transfer learning.**

Keywords- Deep Learning, Convolutional Neural Network, Breast Carcinoma, Transfer learning

I. INTRODUCTION

Our body has millions of cells. The unexpected and uncontrolled production of these cells results in the formation of tissues in the body called tumors. Tumors grow uncontrollably and expand to other regions in the body. According to Globocan data 2022 [1], Breast malignancy is the largest source of fatality in women. Breast carcinoma cases are increasing at an alarming rate, particularly affecting younger age groups. By 2040, it is estimated that there will be 28.4 million new instances of carcinoma and 16.3 million cancer deaths worldwide, leading to a rise of 47% from 2020[1]. Thus, there is an urgent need to diagnose and categorize it at the initial stage to decrease the mortality count.

Interpretations of medical images require trained experts, doctors, radiologists, etc. [2]. Moreover, tumor localization is itself a tedious and time-consuming task. In recent years, CAD, i.e., Computer-Aided Diagnosis, has emerged as a diagnostic tool that aids in analyzing medical images. The approach for evaluating medical pictures using computers originated in the 1960s [3]. Medical image processing using CAD has shown to be a valuable tool for early detection, segmentation, diagnosis, and categorization.

Deep Learning (DL) has been identified as a prominent research area, having applications in broad domains. It is a further enhancement in artificial neural networks (ANNs), with several layers of artificial neurons that improve predicting ability of data. The key benefit of DL approaches is the ability to automatically retrieve significant data representations from the inputs [2].

Convolutional Neural networks (CNN) are among the most effective learning algorithms for analyzing pictures and have shown outstanding performance in detection, segmentation, and classification. CNN could be trained either from scratch or by a transfer learning approach. CNN's trained from scratch don't utilize any pre-trained architecture and consume many resources during training. They require high computing GPUs in the training stages. The transfer learning approach, is employed when the training data from the target destination is significantly less; thus, the expertise gained from the source discipline is transferred to the target discipline. The Fine-Tuning strategy is adopted, and therefore the network is fine-tuned layer-wise. In the current study, we have proposed a model for diagnosing and classifying breast cancer by implementing MSTLA. Three mammography datasets were used, i.e. MIAS (Mammography Image Analysis Society), In-Breast, and CBIS-DDSM (Curated Breast Imaging Subset of Digital Database for Screening Mammography).

The paper is organized in the following way: Section 2 discusses the literature for diagnosing breast malignancy using deep learning approaches. The Proposed method and the datasets are discussed in Section 3, and the Results in Section 4. The conclusion, along with the future scope, is addressed in Section 5.

II. RELATED WORK

In [4], the authors proposed a multi-stage framework for breast carcinoma diagnosis using deep convolutional neural networks(DCNN). They utilized mammography and tomosynthesis data, and the results proved that the AUC improved by a factor of 0.6 by implementing an intermediate stage of the transfer learning approach. The researchers suggested hyper-parameter optimization in the future to further enhance the accuracy and AUC values.

In another research [5], the authors have shown that DCNN based approach outperforms the feature-based approach for mass detection in Digital Breast Tomosynthesis (DBT) images and thus improves the sensitivity from 83% to 91%. The initial training was performed on a mammography dataset, and then the model was further utilized to train DBT images. The AUC value improves from 0.81 to 0.91 in the second stage. In [6], the researchers have proposed a multi-

ISBN: 979-8-3503-9648-5/23/$31.00 ©2023 IEEE

A Framework for Generalization Error Evaluation in Deep Convolutional Neural Networks

Gunjan Chugh
Dept. of Computer Science & Engg.
Delhi Technological University
New Delhi, India
chugh.gunjan8917@gmail.com

Shailender Kumar
Dept. of Computer Science & Engg.
Delhi Technological University
New Delhi, India
shailenderkumar@dce.ac.in

Nanhay Singh
Dept. of Computer Science & Engg.
Netaji Subhas University of Technology (East Campus), New Delhi, India
nsingh1973@gmail.com

***Abstract* -Deep learning has introduced various paradigms in the healthcare industry. Deep Convolutional Neural Network models assist doctors in diagnosis, surgery, and other areas. Incidences of breast tumors are growing at a frightening pace. Thus, it is necessary to diagnose it so that the mortality count can be decreased. Generalizability defines how well the model performs on unseen data. When capturing mammograms different types of noise get added to the images. Noise may significantly diminish classification ability and make class separation more difficult. Thus, analyzing the model's generalization on noisy or unseen data is very important. This paper proposes an approach for analyzing generalization errors in Deep Convolutional Neural Networks by inducting noises such as Gaussian, Salt and pepper, and Speckle. We have utilized the CBIS-DDSM dataset. Three prominent deep neural network models- Inception v3, Dense Net 201, and EfficientNetB4- are fine-tuned to assess the model's efficiency on noisy data. Results proved that the DenseNet201 has minimum Generalization errors i.e. 0.12 and performs pretty well on noisy data with a minimum loss rate. On the other hand, maximum distortion is caused by Speckle Noise in Inceptionv3 and Efficient NetB4 leading to a considerable drop in accuracy. Inception v3 has the highest Generalization Error i.e. 0.61 and thus exhibits minimum generalization capability.**

Keywords: Deep Convolutional Neural Networks, Generalization Error, Medical Imaging, Noise, Breast Carcinoma

I. INTRODUCTION

Cancer has become a significant global health issue all over the world. In 2020, approximately 2,261,419 new instances of breast carcinoma were discovered in women. The World Health Organisation (WHO) has released a new Global Breast Cancer Initiative Framework, which includes saving the lives of 2.5 million women from breast tumors by 2040. Breast Carcinoma has surpassed lung cancer and is the leading cause of cancer death in women [1], [2]. In India, the mortality rate is very high because detection takes place very late[3]. Early cancer detection and diagnosis will boost survival rates and reduce false positive rates. Several imaging modalities, such as mammography, computerized tomography (CT), MRI (Magnetic Resonance Imaging), and ultrasound, can be used to determine breast carcinoma[4].

A lot of research has been performed on diagnosing breast carcinoma using machine learning and deep learning strategies [4]–[8]. These studies have proposed several methodologies for early diagnosis of Breast Carcinoma and have attained pretty good accuracies. Very few research in the literature focuses on the generalizability of Deep Neural Networks. The inability of neural networks to generalize across domains is a significant barrier to their use in improving medical scenarios [9].

Due to their large learning capacity, Deep Convolutional Neural Networks(DCNN) can memorize the training data. Generalization is an approach to analyze how the model behaves on unseen data. Generalization Error(GE) measures the difference between training and testing errors. Leading causes of GE include memorization of training data, overfitting, a model with too many parameters, etc. When Deep Neural Networks are trained on medical imaging datasets, it is crucial to study the effect of generalization error on transfer learning networks[10]. The performance of any model depends upon the generalization ability. Regularization techniques such as data augmentation, weight decay, and dropout reduce the model complexity and thus improve the generalization capability of DNN by preventing overfitting and thus leading to low GE.

Further, the robustness of any deep neural network can also be examined through various Adversarial Attacks. Adversarial attacks represent any perturbations being added to the dataset to befool a deep learning model. Deep learning models used in medical imaging are more vulnerable to adversarial attacks as compared to non-medical DL models[11]. Thus, there is an urgent need to study and analyze the impact of these attacks on our models.

This research presents a novel technique to assess the Generalization error through various Noise samples. The noise affecting Mammogram images includes Gaussian, Salt & Pepper, and Speckle. We have analyzed and plotted the model behavior concerning the various perturbations added to the dataset to analyze the generalization capability of the DCNN.

The paper is organized as follows: Section 2 addresses the literature work in this field. The proposed framework is analyzed in Section 3. Section 4 examines the findings and outcomes followed by the discussion section. Section 5 outlines the conclusion and the future perspectives of the research.

II. RELATED WORK

In[10], the authors analyzed generalization on DCNN by studying approximation error on breast mammograms. They utilized transfer learning using Alex Net and Google Net. Results were examined after corrupting labels and shuffling pixel values. As DCNNs are prone to overfitting and memorization, it's crucial to design a proper transfer learning strategy to reduce generalization errors in DCNNs.

2023 IEEE Engineering Informatics | 979-8-3503-3852-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/IEEECONF58110.2023.10520375

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahabad Daulatpur, Main Bawana Road, Delhi-110042

PLAGIARISM VERIFICATION

Title of the Thesis: **Medical Image Analysis for Diagnosing and Classifying Breast Cancer Using Deep Learning**

Total Pages: **135** Name of the Scholar: **Gunjan Chugh**

Supervisor(s):

(1.) Prof. Shailender Kumar, DTU, New Delhi

(2.) Prof. Nanhay Singh, NSUT (East Campus), New Delhi

Department: **Computer Science and Engineering**

This is to report that the above thesis was scanned for similarity detection. The process and outcome are given below:

Software used: **Turnitin** Similarity Index: **8%**

Total Word Count: **29760**

Date: _____

Candidate's Signature Signature of Supervisor(s)

BRIEF PROFILE

Gunjan Chugh is currently working as an Assistant Professor at Maharaja Agrasen Institute of Technology, New Delhi. She has submitted her thesis entitled "Medical Image Analysis for Diagnosing and Classifying Breast Cancer Using Deep Learning" for the Doctorate Degree at Delhi Technological University, New Delhi. She has more than 11 years of Teaching Experience. She received her M. Tech degree with distinction in Computer Science from Banasthali University, Rajasthan, in 2013 and her B. Tech degree with distinction in Computer Science and Engineering from Maharshi Dayanand University in 2011. She qualified GATE exam in 2011 and UGC-NET in 2019. She has published many papers in reputed journals and conferences. Her two papers are published in the SCIE Indexed Journal with very good Impact factors. She has also published and has been granted a patent in the year 2022. She also serves as a reviewer of many SCI and Scopus Indexed Journals. Her research interests include Artificial Intelligence, Machine Learning, Deep Learning, and Image Processing.

Google Scholar: https://scholar.google.com/citations?user=fa4sCVcAAAAJ&hl=en

ORCID Profile: https://orcid.org/0009-0000-1956-2685