Breast Cancer Detection Using Machine Learning and Transfer Learning

By

Assad Muhammad Rafique

(Registration No: 359382)

Supervised By

Dr. Muhammad Osama Ali

Department of Mechatronics Engineering

College of Electrical & Mechanical Engineering

National University of Sciences & Technology (NUST)

Islamabad, Pakistan

(2024)

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis by Mr. Assad Muhammad Rafique Registration No. 00000359382, of Electrical and Mechanical Engineering College has been vetted by undersigned, found complete in all respects as per NUST Statues/Regulations, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature: OsamaAli

Name of Supervisor: Dr. Muhammad Osama Ali

Dated: 23 Aug 2024

Signature of HOD:

Dr. Hamid Jabbar

Date: 23 Aug 2024

Signature of Dean: Rashid

Brig Dr. Nasir Rashid

Date: 23 Aug 2024

.

DEDICATION

This work is dedicated to my parents and siblings. Their extraordinary support and encouragement have been instrumental in guiding me to this significant accomplishment.

ACKNOWLEDGEMENTS

I begin by expressing profound gratitude to Allah Subhanahu Wa Ta'ala for His divine guidance and blessings throughout this endeavor. My deepest appreciation goes to my parents for their unwavering support and sacrifices, and to my family for their constant encouragement.

I would like to express my sincere gratitude to my supervisor, Dr. Muhammad Osama Ali, for his invaluable guidance and support throughout this research project. His expertise, insightful feedback, and encouragement were instrumental in the completion of this work.

Lastly, I'm deeply appreciative of my friends for their unwavering companionship and encouragement, which kept my motivation intact throughout this challenging yet rewarding journey.

Abstract

This study explored the application of deep learning techniques to enhance the detection of breast cancer. The research aimed to improve the accuracy, reliability, and efficiency of breast cancer diagnosis methods by addressing the limitations of traditional imaging and computer-aided diagnostic (CAD) systems. The study focused on the development and evaluation of advanced deep learning models, including YOLOv8 for object detection and Mask R-CNN for segmentation and tumour size prediction.

The findings of the research indicate that the Random Forest model demonstrated the highest accuracy in identifying various BI-RADS categories, supporting the reliability and effectiveness of the model in breast cancer detection. The integration of these advanced deep learning models into the clinical workflow can streamline the diagnostic process, reduce false positives and negatives, and improve patient outcomes by enabling early detection and treatment.

The study contributes to the field of breast cancer detection by showcasing the transformative impact of deep learning in addressing the complex challenges of medical imaging. The comparative analysis of different models provides valuable insights and a foundation for future research and development efforts in this area.

Keywords: Breast cancer detection, Deep learning, Random Forest, Medical imaging, Clinical diagnosis

Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Research Background

Breast cancer, the most common type of cancer, has become a leading cause of death among women across the globe [1]. The “World Health Organization” reports that breast cancer accounts for about 2.3 million annual cases, making it the most frequent cancer type in both men and women combined. Breast cancer is the primary or secondary cause of cancer-related deaths in women in 95% of the world's nations. Breast cancer is becoming more common and takes more lives each year. It is anticipated that by 2040, there would be one million breast cancer-related fatalities and over three million cases of breast cancer worldwide each year [2]. It is argued that breast cancer screening enhances the chances of early detection and hence, results in better outcomes and survival rates [3]. Mammography, ultrasound, and “Magnetic Resonance Imaging (MRI)” have been employed as screening modalities for breast abnormalities from the past few years [4]. However, these methods are not always accurate and can show false negatives and false positives, which may lead to more biopsies, stress, or failing to diagnose a cancer at an early stage. [5].

Nowadays, “Deep Learning (DL)”, that is a part of “Artificial intelligence (AI)” and based on how the human brain works while recognizing data and making decisions, has impacted almost every field including medical imaging [6]. “Convolutional Neural Networks (CNNs)” have also performed excellently in other image classification tasks, for which they can be used in the analysis of medical images [7]. DL has been effectively applied in the detection of breast cancer because it is able to detect and learn small features of images making accurate diagnosis of breast cancer possible [8]. Some of the DL methods applied in the diagnosis of breast cancer involve the use of CNNs on large datasets of mammogram images that have been labelled. These networks are trained to identify characteristics and structures of cancerous and non-cancerous tissues [9]. Albawi, Arif [10] showed that DL algorithms could be used to diagnose even the skin cancer from images with the same accuracy as dermatologists. This success has prompted the use of same techniques in breast cancer detection.

One benefit of DL in the detection of breast cancer is that it can handle heterogeneity of cancer well. Breast cancer has various types and phases and, therefore, the likelihood of achieving high accuracy with standard methods is negligible. DL models, nevertheless, can be trained on such diverse dataset that can include different types of cancer, stages, and patients. It also helps the models to learn well and be consistent while working in various clinical settings [11]. Also, DL is useful for radiologists as it reduces the workload, and the time required to make the diagnosis. Most radiologists interpret many mammograms, and this leads to fatigue which in turn, contributes to error. The first level of screening could be done using DL-based algorithms and the radiologists could then work on those cases which they can handle best. When implemented as a helper, DL models are useful to radiologists and improve the detection rates by decreasing the likelihood of missing a diagnosis [12]. Figure 1.1 shows the applications of the machine and deep learning techniques in the detection of breast cancer:

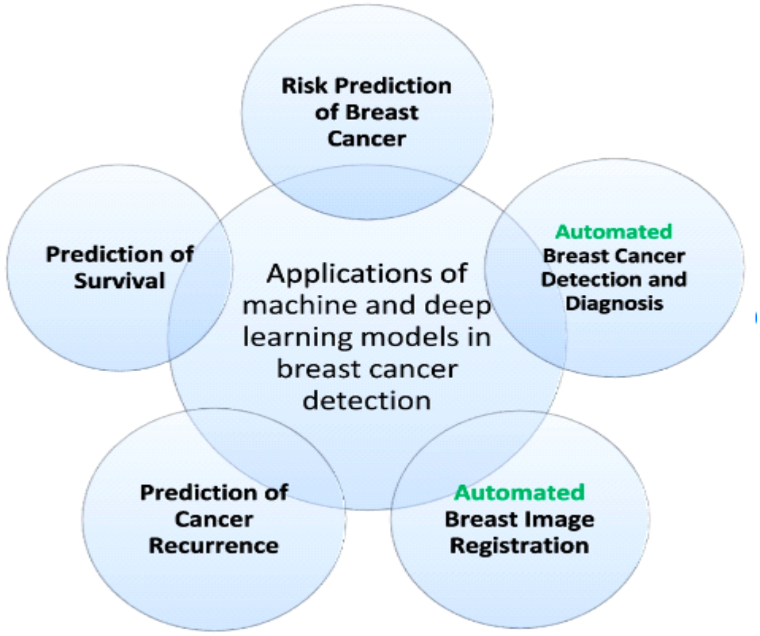

Figure 1.1. Applications of the Machine and Deep Learning Techniques in the Detection of Breast Cancer

(Source: Neha Thakur [13])

Nevertheless, there are several limitations as well that have to be considered to apply DL in breast cancer detection autonomously. There are not enough big datasets that are annotated and, at the same time, of good quality. DL models require high amounts of training data which need to be labelled; however, such datasets are scarce mainly because of privacy concerns and the time-consuming process of data labelling [14]. It is necessary to build large and diverse samples with the help of cooperation between research institutions, hospitals, or other medical organizations [7]. The other disadvantage of DL models is that they are not easy to interpret. CNNs have high accuracy; however, the decision-making process within the algorithm is not easily explicable and is sometimes called the "black box". Such lack of transparency can harm the adoption of AI systems in clinical practice [15]. To improve the interpretability of the models, work is being done on how to represent the learned features in the form of visualization or through the use of attention mechanisms that highlight which parts of the image are useful [16].

However, applying DL models in real clinical environment requires proper validation and normalization procedures. A model trained on some data may not be ideal for another population or imaging modality. Therefore, the external test and validation of these models across multiple clinical centres and in different patient groups are crucial for their generalizability. The ethical perspectives also need the nod from the regulatory authorities and ethical mandates to safeguard the patient and their information [17]. In addition to the above challenges, there are some general ethical concerns regarding AI application in healthcare. Problem that rises with DL is the algorithmic bias. In other words, the accuracy of the models trained using non-diverse data may be significantly worse for particular types of patients. Such aspect can be handled by having diverse images in the training dataset and using bias mitigation procedures. Moreover, when implementing AI systems, it is necessary to support clinical decision-making while leaving the control to clinicians [18]. Nevertheless, there are many possibilities of using DL for breast cancer diagnosis. This is done to improve algorithms in detecting these diseases and secondly, to make algorithms used in clinical practice more reliable. The future of DL for breast cancer detection entails collaboration between researchers, doctors, and regulatory bodies [19]. Today's advances in this area have stemmed from the availability of computing resources and big data. Other solutions have also been proposed, such as transfer learning, where models learn from large sets of images and are then further adapted for specific applications, which can also be useful to counter the small size of many medical datasets [20]. For example, by combining several types of data such as incorporating patient history and genetics on top of the mammography data, DL models' accuracy and resilience can be enhanced [21].

The application of DL for breast cancer detection is promising and has the potential to be used in other forms as well. Uses of DL models include treatment response prediction, disease prognosis, and developing customized treatment plans based on the patient's characteristics. These models can be retrained from new data and clinical feedback and improve the capacity to deliver more precise and personalized breast cancer care routinely [22]. Another valuable approach is the use of AI in multimodal imaging where information from different imaging modalities (for example, mammography, ultrasound, MRI) is combined to improve accuracy of the diagnosis. It is also important to note that in the multimodal DL models, the cross-modal information is helpful for better detection and further characterization of the breast cancer [23]. This potential of DL is even more profound in breast cancer detection due to its nature as it has the ability to learn from new data. From the availability of more annotated datasets and more technology implementations on DL, the model's performance is expected to enhance [19]. Moreover, implementing AI methods in the telemedicine and mobile health systems can make breast cancer screening more accessible for women who struggle to reach the radiologists, specifically in the developing countries [24].

1.2. Thesis Motivation

Breast cancer remains a leading global public health concern and its detection using DL has received much attention since the past few years [25]. Nevertheless, there are several research gaps in the literature which limit the applicability of DL in this vital area. Among all of these gaps, the one that can attract the most attention is the data accessibility and compliance. The deep learning models require a vast amount of high-quality labelled data to achieve passable results [26]. Generally small and heterogeneous databases are available to the public for breast cancer detection [11]. Prominent examples are the "Digital Database for Screening Mammography" and the "Breast Cancer Digital Repository"; they often provide a relatively small number of images compared to other branches of machine vision [27]. Moreover, these datasets are mostly less annotated, and the significant parts are missing, such as tumour edge or much pathologic information which is crucially important for the successful training of a proper DL model. This lack of large-scale datasets hampers the prospect of creating universally relevant and strong DL models across populations and imaging modalities [28].

Another important concern is what authors define as data heterogeneity. Breast cancer images can come from various modalities like mammography, ultrasound, MRI, and histopathology [29]. Though each of these modalities provides complementary and diverse information, the integration of such heterogeneous sources is the primary focus of the recent research [30]. Most of the studies, such as the ones by Adam, Dell'Aquila [31] and Loizidou, Elia [32], are based on a single type of imaging which limits the models to use all the potentially diagnostic modalities. For example, there is the possibility of improving the levels of detection accuracy by incorporating data from more than one modality but multi-modal deep learning approaches have not been pursued extensively [33]. Additionally, variations in image acquisition methods, settings of the equipment, and patients can complicate efforts to measure patterns that can be generalized [34]. Another emerging issue relates to interpretability of DL based models. In fact, DL models, particularly CNNs, perform well when used in breast cancer diagnosis; however, due to the complexity of these models, clinicians are reluctant to trust these systems and apply them practically. This lack of decision making presents a serious challenge to clinical integration [35]. In the literature, there are only a few studies, like the one by Vázquez-Lema, Mosqueira-Rey [36], that try to develop DL models that can explain what about the images makes them detect the necessary features and areas. It is crucial to enhance the DL models' interpretability so that DL models can be well implemented and applied in clinics [37].

Further, there are no assessment and validation practices leading to DL models. Some studies show high classification accuracy on either internal or on the limited datasets that are publicly available but for research purposes only [38]. The use of DL models, however, depends on the population investigated, the imaging machines used, or some clinical context [6]. But, validation of these approaches across different institutions and creating a longitudinal database of patient outcomes are necessary to assess the usefulness of such measures. Furthermore, the absence of a standard makes it difficult for DL to have well defined and readily comparable values and thus, hampers progress in the field [39]. Another topic that should be studied further is the application of DL models in the clinical practice. Most research work, such as the one by Zhai, Yousef [40], focuses primarily on identification and validation of individual models instead of describing how such systems might be incorporated into operations. When it is implemented in a clinical environment, some of the challenges that are likely to be encountered include; how to interface the DL models, how to make it work in real-time, and how to integrate it with existing healthcare systems [41]. Moreover, there is uncertainty about the feasibility and benefits of adopting decision-support tools based on DL for enhancing diagnostic accuracy, time, and patient satisfaction in clinical practice. For instance, the practical concerns need to be met to facilitate the proper application of DL models within clinical settings [42].

Biasness in DL models is another research gap that has drawn a lot of controversy. Conventional DL models also have the ability to incorporate behavioural bias into the model and this may be reflected in the results that might vary based on demography. For example, models trained predominantly based on data from one population may predispose very different predictions when applied to other population due to genetic discrepancies or different exposures and eating habits [43]. Making DL models fair and unbiased entails the use of training and evaluation sets that include a diverse representation of features with regard to sensitive attributes and the capacity to identify and address bias during model development [44]. But there is not enough work done on investigating the bias and fairness of DL models for breast cancer detection.

In addition, there is a dearth of literature regarding the combination of DL with other advanced technologies, including radiomics and genomics, in breast cancer diagnosis. Radiomics can be described as the process of obtaining qualitative features from medical images that may be advantageous in terms of tumour profile [45]. The radiomic features that can be derived from MRI can be combined with DL models to improve the overall performance of the diagnostic and prognostic capabilities of the models [46]. Similarly, integrating genomics with additional imaging and DL could lead to more precise detection and treatment of breast cancer [47]. However, the integration of these diverse data and the formation of multi-disciplinary practices are not fully discovered.

Besides, the potential for ethical or legal discussion of DL in relation to breast cancer screening is neglected in the literature. While discussing the issues that relate to the use of DL models in the sphere of healthcare, one should mention privacy, informed consent, and liability upon an incorrect diagnosis [48]. For these issues and for the purpose of having proper ethic and law compliance in DL applicability in breast cancer detection, there is a need of guidelines [49]. In addition, patients, clinicians, and other stakeholders should participate in DL systems' development actively by expressing their vision and expectations [50]. Therefore, further research is required in the area of detecting breast cancer using deep learning to fill gaps that exist and to develop the full potential of the technology. This is because the current approaches in data availability, data quality, and data heterogeneity pose challenges in the formation of more refined and generally applicable models [51]. The use of such datasets containing different population and image types has a high potential to enhance the accuracy of DL models. Furthermore, increasing the interpretability of these models is crucial for clinical practice since clinicians rely on the models to make decisions on the patient's behalf [52]. Performance evaluation of DL models should thus, undergo rigorous assessment and verification to ensure they perform to the same level in other clinical settings and with other patients [53]. The application of DL models in clinical practices raises issues such as the layout of the user interface, real time execution of the models, and compatibility with existing systems [41]. Therefore, conducting this research will help enhance the breast cancer detection by improving the accuracy of the AI systems used in medical diagnosis.

1.3. Justification and Rationale

The reason why deep learning should be used in diagnosis of breast cancer is because it offers more accurate results compared to the traditional techniques [54]. DL can interpret the mammogram, ultrasound, or MRI scans and diagnose the images with higher accuracy and precision than the radiologists. It can detect small variations on tissues that a radiologist might not notice and this means that chances of false negatives are rare and more breast cancer cases are detected at early stages when they are easier to manage [55]. Moreover, the existing deep learning models can provide quantifiable results like tumour mass and proliferation rate, which are beneficial for diagnosis and monitoring [56]. Several more straightforward advantages of deep learning in breast cancer detection include the potential to decrease the burden on professionals [12]. Diagnostic imaging is also overused by radiologists and oncologists across all patients, resulting in high stress levels and a higher possibility of errors [57]. Therefore, deep learning systems work in conjunction with human healthcare professionals by filtering out myriad unremarkable cases, while identifying those that require human assessment. This not only improves the efficiency of the diagnostic activities but also the quality of their treatment [12].

Moreover, the implication of DL in the detection of the disease is an indication that disparities in healthcare services may be minimized [58]. There are not enough qualified radiologists and oncologists in rural areas and in the countries with low income [59]. These regions can then be able to use deep learning-based diagnostic tools to provide timely and accurate diagnosis of breast cancer without the need for doctors. Thus, such a democratization of diagnostic services may eventually lead to higher rates of early diagnostics and better health outcomes for the patients who otherwise would not have access to adequate quality healthcare [60]. In addition, the continuing advancement of DL technologies as well as the improvement of the neural network structure and the accessibility of even larger and more diverse data sets improve the efficiency and reliability of these systems [61]. More scientific studies and cooperation between computer scientists, medical researchers, and healthcare practitioners are necessary for improving these technologies and implementing them in medical practice [62].

In the literature, much emphasis has been laid on the need to further explore the use of deep learning to enhance diagnosis of breast cancer as it holds great promise in changing the current diagnostic approaches. For example, in one of the recent reviews, Najjar [6] note that the use of the AI and DL models is as effective or even more effective than the use of radiology in the screening of breast cancer and he urged for further research for enhancing the availability and application of these technologies. In another study, Huang, Wei [63] reveal the effectiveness of incorporating AI algorithms in reducing the workload of radiologists. They indicate that further studies should be conducted to enhance the rate at which DL models work and help the radiologist, especially in centres that attend to several patients. In addition, Castiglioni, Rundo [64] offer a comprehensive overview of deep learning for medical imaging and state that additional investigations are required to determine the effectiveness of such algorithms in real-world contexts. The authors point out that large datasets are required in deep learning to train and test the model to increase its applicability across different populations. Altogether, these studies emphasised the need for further research to provide better understanding of the application of deep learning in the detection of breast cancer.

1.4. The Aim and Objectives of the Study

Therefore, the aim of the study is to enhance the breast cancer detection using deep learning techniques. In order to fulfil the aim, the following are the research objectives:

- To develop a deep learning model that will be able to provide accurate detection of breast cancer.
- To enhance the detection and treatment of breast cancer at an early stage in order to increase the survival rates.
- To reduce false negatives and false positives in breast cancer detection to give out more accurate diagnosis.
- To establish the model's credibility and reliability by conducting extensive experiments and verifications on the data set.
- To promote the adoption and application of the developed deep learning model into clinical works.

1.5. Significance of the Study

The study will impact the medical diagnosis, patient outcomes, health care choices and decision, and access to quality health services. Breast cancer is among the leading causes of death in women globally and as the disease in most cases is not easily noticeable during early stages, it is wise for women to undergo screening so that the disease can be noticed at an early stage [65]. The general goal of this study is therefore, to mitigate the shortcomings of the traditional diagnosis procedures by applying deep learning and be one of the strides in the fight against breast cancer. The study is helpful as it can contribute to the increase in the diagnostic reliability. Mammography, ultrasound and MRI are the common modalities used in the detection of breast cancer and all these modalities necessitate the involvement of a radiologist. However, these techniques are not completely accurate and might create false positive and negative results with consequent anxiety, invasive investigations, or delayed treatment [66]. DL algorithms with sufficient data can be trained and tested to deliver high accuracy in analysing medical images and sometimes, even surpass human observer performance. It can also reduce the number of diagnostic errors carried out, thereby allowing a larger number of patients to be diagnosed correctly and as early as possible. High levels of accuracy in deep learning models are also useful for identifying pre-stages that are easier to treat, enhancing patient experiences and longevity [67].

One more advantage of this study is that it has the potential of easing the load of work among the health care personnel. For example, radiologists and oncologists should be capable of analysing multiple diagnostic images, though this ultimately causes fatigue and can result in mistakes. Consequently, with the use of deep learning algorithms it is possible to reduce the time spent on this step. Not only does this enhance the diagnostics process, but it reduces the burden on clinicians and staff overall, thereby increasing overall satisfaction and reducing burnout [68].

Besides, the study has the significant utility from the perspective of increasing the chances of identifying breast cancer. At present, in many areas of the world, especially the developing countries, there is a shortage of skilled radiologists as well as advanced imaging equipment. Similarly, patients in these areas are offered inferior medical services, and often times are diagnosed later than patients who seek care from better-endowed institutions [69]. Once the deep learning diagnostic tools are set, they can be used in such regions with effective diagnostic ability of breast cancer without reference to the medical professionals. This democratization of healthcare could eliminate the geographical as well as socio-economic cancer care inequalities since all female individuals in various regions and wealth brackets could receive accurate screening for early breast cancer [70].

This study also contributes to the practical research in the field of healthcare advancement and technologies. The subject of diagnostics in medicine has now been augmented by the incorporation of artificial intelligence [71]. Given that this present study gives evidence about the efficiency of deep learning in identifying breast cancer, the knowledge can be utilized in other fields of medicine as well such as in diagnosing lung cancer, prostate cancer, or any other kind of cancer. This also provides a new approach to technology advancement when performing research that requires collaboration from computer and software engineers, medical and healthcare professionals. In addition, the study can contribute to the development of the general knowledge about the use of AI in the healthcare domain, generating interest in its further exploration. Therefore, the results of this study can be used to create a higher level of the architecture, the development of deep learning technologies, and the increase in the amount and the varieties of the data pool which may also lead to multiple improvements in diagnostic accuracy and reliability in the future. The demonstration of the deep learning models for practice in the clinical setting can also assist in building faith among the practitioners and patients on the need to incorporate the AI tools as standard practices in the system [72]. Finally, the study has major implications for the advancement of the concept of the personalized medicine. Therefore, through the use of deep learning algorithms, it is possible to develop an individually tailored treatment plan based on the exact features of each tumour type in the patient. Such personalizing in cancer care can enhance the results of the treatments, reduce side effects, and thus, have a favourable effect on the quality of life in patients. The conclusions derived from enhanced imaging study are also beneficial in clinical management of the patients since changes in disease process and response to therapy can be monitored [73].

1.6. Structure of the Study

This study has six chapters. Chapter one introduces the breast cancer detection problem, the significance of early detection, and the opportunity for utilizing deep learning to improve existing solutions. Chapter two reviews the previous studies and compares the previous conventional methods and deep learning works done for breast cancer detection to identify the knowledge gap for further research. Chapter three outlines the research method, data acquisition process, data pre-processing, and architecture of the deep learning models in the work. Chapter four, which is about experimentation, describes the experimental procedures such as the training processes and the evaluation criteria. Chapter five presents the results of the experiments and the discussion regarding the applicability to the clinic. Chapter six concludes the study by outlining the research findings, the research limitations, as well as future research directions to enhance breast cancer detection using deep learning

Chapter 2: Literature Review

2.1. Introduction

Cancer is one of the major worldwide health issues; in 2020, 10 million people died because of its worldwide [74]. Attributing one in six fatalities worldwide, it is the second most common cause of death. The likelihood of survival for almost all malignancies rises dramatically with early detection, diagnosis, and treatment [75]. One of the main causes of mortality for women globally is breast cancer. Therefore, the likelihood of recovery and the death rate are increased with early diagnosis and identification. In accordance with the World Health Organization, globally, 626,700 women die every year from cancer-related causes, and of these, the most common diagnosis worldwide is breast cancer [76]. But, if it is found early on, treatment expenses can be cut significantly, and the death rate can be significantly decreased.

2.2. Overview of Breast Cancer

Cancer cells exhibit abnormal behaviour that goes beyond the usual life and death cycle. Self-sufficiency and self-management may be an evolutionary process in cell division. Cancer cells advance by selected transition to malignancy, similar to how organisms evolve through natural selection and variation [77]. The most prevalent disease in women globally, breast cancer (BC) is second only to lung cancer in industrialized nations in terms of cancer-related deaths. The death rate has decreased by almost close to forty percent while the yearly incidence of female BC is still rising and is mostly caused by hormone receptor-positive, no metastatic ailments [78]. The hormonal receptor status and phase at diagnosis play a major role in the prognosis of breast cancer. The prognosis for early-stage breast cancer is usually beneficial, with a nearly 100 percent 5-year survival rate. Advanced stages of breast cancer significantly lower the 5-year survival rates, with the chances for stages II, III, and IV standing at about ninety-three, seventy-two, and twenty-two percent respectively. All instances of breast cancer have the highest 5-year survival rates when they are positive for estrogen and progesterone receptor, and the lowest when they are negative for estrogen and progesterone receptor [79]. Reproductive and non-reproductive risk indicators are categorized as important variables, and they are all impacted by economic growth. The risk of breast cancer rises with age at menarche, age at menopause, number of children, and duration of breastfeeding. Higher human development tends to reduce the average age at menarche because it improves average nutritional status, which is a major factor in determining the age at which menarche begins. The two non-reproductive risk factors for breast cancer that are of particular interest are increased alcohol consumption, which is estimated to account for four percent of all cases of breast cancer diagnosed, and obesity, which has been shown to double the risk of breast cancer in overweight post-menopausal females. A genetic or inherited factor, accounts for five to ten percent of breast cancer cases [80]. Breast cancer is treated with medical procedures, chemotherapy, radiation, hormonal treatment, targeted treatment, and immunotherapy. Several specialist physicians must work together to complete the therapeutic regimens. The usual course of treatment for non-metastatic breast cancer is surgery, and preoperative systemic treatment based on treatment with chemotherapy can lower the tumor size of the breast. Surgery is only utilized as palliative therapy in a small number of metastatic patients, and systemic treatment is still the chosen course of action for metastatic breast cancer [81].

In addition to having a higher chance of survival, those with cancer who receive an earlier diagnosis also tend to receive better care, suffer fewer side effects from therapy, and have higher quality of life than those who receive a later diagnosis. Complex and multidimensional, efforts to enhance early cancer detection have been at the forefront of worldwide governmental and charitable (such as Cancer Research UK) projects. An earlier cancer diagnosis can be aided by two different patient actions. These include going to cancer screenings, which are designed to find cancer before symptoms appear (for example, mammograms for breast cancer), and showing up early for primary care visits when they think they may have cancer [82].

2.3. Conventional CAD (Computer-aided Detection) Systems

The early detection of breast cancer is greatly aided by the processing and interpretation of breast radiography pictures [83]. Radiologists should be able to visually identify any spot that does not appear to be normal tissue since this might be a direct sign of cancer risk. Experts frequently look for patches of white during the diagnostic process, noting their location inside the breast as well as the thickness of the fat tissue and/or the size and form of the breast. A malignant tumor has the potential to enlarge and transform into a mass lesion. A benign tumor, on the other hand, is typically not malignant, presents no health concern, and is not expected to spread. White patches and dots are another appearance of uneven tissue that is often benign. However, to monitor their form and pattern, which might indicate a cancerous indication, ongoing follow-up is necessary [84]. According to Guo, Xie [85], computer-aided diagnostic CAD employs algorithms for pattern recognition to identify foreign shapes in an image for a doctor to consider when diagnosing a patient. CAD systems use many methods of imaging, including magnetic resonance imaging, ultrasound, computerized tomographic, mammogram, and biopsy, to diagnose breast cancer. By eliminating the need for lengthy readings and maintaining the consistency of lesion detection, CAD also improves radiologists' analytical routine and interpretation proficiency. Using CAD, machine learning has been used in breast imaging. Computer-aided detection (CAD) systems have demonstrated encouraging results in breast imaging as a tool to assist radiologists in reading screening mammograms, they may generally result in a rise in false-positive recalls that radiologists must reject [86].

Ramadan [87] stated that medical image analysis and interpretation, such as determining the precise position and probability of carcinoma in a suspected lesion, are aided by computer-aided diagnostic (CAD) systems. To distinguish between two strands of CAD systems, utilize CADe and CADx schemes. A key difference between the two is that while CADe refers to computer-aided detection systems, ascribed to the picture to aid in the discovery and singling out of possible anomalies, it shifts responsibility for interpretation back to radiologists. Conversely, CADx refers to the computer-aided diagnosis, wherein radiologists could use a CADx to aid them in making decisions by characterizing results from radiological images recognized either by a CADe system or a radiologist [87]. According to Baccouche [84], the integration of AI with computer vision has resulted in workable computer-aided diagnostic

(CAD) systems that help physicians with diagnosis tasks. An automated computer-aided detection system (CAD) might be useful as a medical imaging application to accelerate the screening process and offer a second reading that helps uncover any undetected trends or indications of cancer that could be missed. The CAD system may also help reduce the number of erroneous positive and negative diagnoses as well as the death rate.

The computer-aided diagnosis (CAD) system is used in medical image processing to assist physicians in making final decisions on various illnesses, particularly malignancies. The goal of the entire procedure is to get important data from medical imaging, including CT, MRI, and ultrasound scans. Numerous computer-aided diagnosis (CAD) systems have been created to diagnose a variety of illnesses, such as lung cancer, breast cancer, and tumors [83]. In mammography, X-ray-based CAD systems are utilized to detect and emphasize micro calcification clusters, microscopic lumps in thick tissue, the marginal structure of suspicious masses, structural aberrations, and the very dense tissue structure. Mammography-based CAD is the ideal method for screening since it can detect irregularities in the breast before indications appear. MRI uses powerful radio and magnetic waves to provide detailed information about inner breast tissues, making it extremely useful for recognizing high-risk BC patients [88].

The initial step in a fully integrated CAD system would be to locate and identify concerning lesions and differentiate between distinct kinds of lesions, such as mass, calcification, architectural distortion, etc. Secondly, the acquired area of interest (ROI) around the breast lesion should be segmented by the CAD system in order to identify its anatomical contour and eliminate its tissue backdrop without compromising the accuracy of its structure. Lastly, pathology-related diagnostic data might be retrieved in order to categories the selected lesion as benign or malignant and to determine its features, like a tumor grading using the Breast Imaging Reporting and Data System (BI-RADS) score or/and a shape classification. A quick and correct judgement must be made by generating each output data exactly since the automated method depends on linked phases [84]. Eltrass and Salama [89] designed a method for identifying breast cancer. The radiographic picture was converted into a binary image as a pre-processing phase, and all areas were then sorted to determine the mammogram's largest area, or the breast region. Pectoral muscle and all artefacts were also removed. For segmentation reasons, this CAD system used the expectation maximization approach. Bratinčević and Matijaš [90] stated that the images acquired in digital format, such as those obtained by DM, DBT, and MRI, may be easily analyzed using CAD algorithms. The effectiveness of the CAD system is evaluated in a number of ways, including laboratory data analysis and an investigation into how CAD affects radiologists in actual clinical settings. The latter two methods yield results that are thought to be the most accurate because they measure the real impact of CAD in clinical practice. According to the Bratinčević and Matijaš [90] study's findings, the initial objective of this novel technique detecting cancers at an earlier stage was not met by CAD when it was used to analyses conventional mammography images, and as a result, the method did not achieve clinical diagnostic productivity. In addition, a decline in the test's specificity was noted, and the challenge of reducing the search's distinctiveness while increasing patients' avoidable recall rate forces one to question whether there really is a benefit to utilizing a CAD system. This is in addition to the system's failure to live up to the initial evaluation stage expectations.

2.4 Breast Cancer Detection Using Deep Learning Methods

Deep learning (DL), a branch of machine learning (ML), Because of its capacity to handle massive volumes of data, has demonstrated exceptional outcomes in a number of disciplines, most notably in the biomedical industry. By effectively analyzing the high multidimensional and connected data, DL approaches automatically extract the features. By using radiographic and histological pictures, the potential and efficacy of models of DL have also been tested and assessed in the detection and prognosis of BC, and they have shown excellent performance [91]. The accuracy of a breast cancer diagnosis can be significantly increased by the quick and efficient identification and diagnosis of tumors utilizing image analysis and machine learning techniques. The diagnosis of clinical diseases, evaluation of treatment outcomes, and identification of abnormalities in various body organs, including the brain, the breast, and the lungs. Medical imaging is the term for specialized methods of examining the human body in order to detect, monitor, or cure a disease [92]. According to Nasser and Yusof [25] a machine learning technique known as deep learning uses learning model to autonomously separate feature representations from input information. In contrast to conventional learning techniques, deep learning may function at its best without the requirement for human-engineered features. Convolutional neural networks (CNN), recurrent neural networks (RNN), deep auto encoders, multi-layer perceptron's and transfer learning and fine-tuning are some of the deep learning techniques that have been proposed in last few years. Many domains, such as processing natural language, algorithms system, image processing and medical imaging have seen the use and success of these methods [25].

CNNs are the most widely used deep learning algorithms that have been proposed in the literature. The two-dimensional input-image structure requires a specific modification to the CNN architecture. A CNN-training process needs a lot of data, which is scarce in the medical field. Using a fine-tuning approach in conjunction with the TL technique using a natural-images dataset, like ImageNet, is one way to solve this issue [93]. Convolutional neural networks are multilayer feed-forward neural networks that employ perceptron's for directed learning and data analysis. They are a type of deep learning method. It is often applied to visual data categorization of images. CNN is mostly utilized in medical imaging to categories and forecast medical data. CNN's design differs from other neural networks' architectures. Tensors or a matrix with more than two dimensions are how CNN displays the image. Arrays inside arrays are arranged to create tensors [94]. According to the results of the study Allugunti [95] indicate that artificial neural networks (ANNs) were initially employed in the field of human image analysis approximately . The algorithms that were most commonly used were PNNs and ANNs. However, morphological and textural characteristics were used in most feature extraction research. It was clear that Deep Convolutional Neural Networks were very helpful in detecting and treating breast cancer early on, which eventually resulted in more effective treatment

2.4.1. Convolutional Neural Networks (CNNs)

Breast cancer is formed in the breast cells which are considered as a common type of cancer in the women population. It is a life-threatening disease, particularly of women after the lung cancer. A method known as convolutional neural network (CNN) was proposed by a study in order to boost the automatic identification of the breast cancer by evaluating hostile ductal carcinoma tissue areas in the whole slide images (WSIs). The study investigated a proposed system that utilizes different CNN architectures to detect the breast cancer automatically comparing the results from the machine learning (ML). The proposed system was successful as it achieved the results with 87% accuracy and could lessen the mistakes made by the humans in the process of diagnosis. In addition, the study found that the proposed system achieved higher accuracy more than 78% of the ML algorithms and it enhances the accuracy by 9% more than the results from the ML algorithms [96].

As breast cancer is the most prevailing type of cancer world-wide and affects millions of women every year. It also leads to the largest number of deaths in the women dying with cancers. During the past few years, the scholars proposed different CNN models to facilitate the process of diagnosis of breast cancer. The CNN are depicting promising results in order to classify the cancers by utilizing the image data sets. Yet, there is a lack of dome standard models that can claim the strong model due to the unavailability of the large datasets which can be utilized for training models as well as validation. A study considered proposed a CNN that outperformed pre-trained models to the various performance measures along with the consideration of eight different fine-tuned models to identify how these models categorize breast cancers applying on the ultrasound images. Thus, the proposed model demonstrated 100% accuracy and achieved 1.0 AUC score; the best pre-trained model represented 92% accuracy along with 0.972 AUC score. The visualization technique of the Grad-CAM heat map represented that proposed model extracted significant features to categorize the breast cancer [97].

Another study developed deep CNN to the segment and classified different types of breast anomalies like the calcifications, carcinomas, masses unlike the available research work that classified the cancer into malignant and benign, contributing to enhanced disease management. It was found that the proposed deep learning model gained a performance of 88% in the categorization of the four types of breast cancer anomalies like calcifications, carcinomas, masses and asymmetry mammograms [98]. Another study found that CNN depicted a better performance in the image classification as compared to the feature-based methodology and represent a promising performance in the medical imaging (MI). CNN was typically employed to classify the lesions as benign or malignant utilizing the MRI images. Utilization of the pixel information and multi-layer CNN architecture with an online data augmentation was structured and later on the CNN architecture was tested. 98.3% accuracy of the network and 0.0167 error rate was observed [99]. Scholars used “Breast Cancer Wisconsin (Diagnostic) Data Set” that contained diagnostic features as well as the clinical features of the breast cancer patients. The paper studied different neural network models such as CNN and the feedforward neural network for the classification of the breast tumors as malignant or benign. According to the results, the CNN gave the highest accuracy of about 98.2% among other selected models [100].

The "deep convolutional neural model network (DCNN)" is utilized for the feature extraction in breast cancer. A known DCNN architecture which is named as AlexNet is utilized and is further fine-tuned for the classification of the two classes instead of 1000 classes. The fully connected (fc) layer is attached to the support vector machine (SVM) in order to gain better accuracy. The outcomes were obtained by using the "digital database for screening mammography (DDSM)" "Curated Breast Imaging Subset" of DDSM (CBIS-DDSM). It was observed that the training on large number of data gave high rate of accuracy. The biomedical dataset contained relatively smaller number of samples due to the limited volume of the patient. The accuracy of new-trained DCNN was 71.01%. highest area under the curve (AUC) obtained was 88% for samples gained from CBIS-DDSM, and the accuracy of DCNN was increased to 73.6% [101].

The CNN was employed in a study to detect the breast abnormalities for the early diagnosis of the breast cancer that was based on the accuracy for the identification of a better method for diagnosis of the breast cancer malignancy. Deep comparison of the functioning of every network as well as its design was performed and then evaluation was done based on accuracy of classification and diagnosis of the breast malignancy through the network in order to decide which specific network outperforms the other network. CNN was found to deliver slightly higher accuracy than other network for the detection and diagnosis of the breast cancer [102].

2.4.2. Recurrent Neural Network (RNNs)

A study aimed to frame the mammogram breast detection model by utilizing the optimal hybrid classifier. Tumor segmentation, image pre-processing, detection and feature detection are the main functional phases of proposed breast cancer selection. Two deep learning architectures were termed as the convolutional neural network (CNN) and the recurrent neural network (RNN). In addition, the grey level co-occurrence as well as gray level run-length matrix was considered as input to RNN and tumor segment binary image was considered as input to the CNN. The outcomes of the study represented that AND operation of the two classifier output tend to give the diagnostic accuracy that outperforms conventional models [103]. Another study used a parallel structure that consisted of CBB as well as RNN for the image feature extraction that is different from common available serial methodology of the extraction of the image features by the CNN and inputting them in RNN. The proposed model of the study achieved the state-of-the-art outcomes on three datasets that demonstrated that the benefits of the parallel structure of the deep neural network utilizing CNN and RNN with the attention mechanism [104]. The Modified Recurrent Neural Network (MRNN) is used for the classification of the breast cancer in the malignant and the benign [105].

Researchers utilized the RNN for the classification that had never been used earlier for the classification of the breast cancer. The performances of the classifiers were compared with the similar implemented before in the respective domain. The outcomes represented that RNN yielded the best performance and the highest accuracy of about 98.49%. though to the best of the scholar's knowledge, the RNN was not implemented in this domain earlier [106].

A study proposed optimized deep RNN model which was based on the keras-Tuner optimization technique and RNN for the diagnosis of breast cancer. Optimized deep RNN comprises of input layer, five dropout layers, five hidden layers and output layers. The three feature selection methodology was utilized to chose significant features from database. The five regular models of ML such as decision tree (DT), K-nearest neighbour algorithm (KNN), random forest (RF), support vector machine (SVM) and naïve Bayes (NV) were contrasted with optimized deep RNN. The RNN as well as the regular ML model were applied the chosen features. The outcomes represented that optimized deep RNN along with selected features by the univariate achieved the most high performance for testing results in comparison to the other models [107]. Researchers focused on the construction of the automated system through the employment of deep learning which is based on RNN models. A GRU-LSTM-BRNN stacked model was proposed in the study which accepts the health records of patient for the determination of the possibility of being influenced by the breast cancer. The proposed model was compared against the other classifiers like the "stacked GRU-RNN model", "staked LSTM-RNN" and "simple-RNN model". The comparative outcomes obtained in the study indicated that "stacked GRU-LSTM-BRNN model" yielded a better classification performance for the predictions linked to the breast cancer [108].

Another study used the classical DT, support vector machine (SVM), linear discriminant (LD), ensemble technique (ET) and logistic regression (LR). RNN and deep neural network (DNN) were utilized for the comparison. The feature selection and its impact on the accuracy was evaluated. The findings showed that only DT and ET showed 98.7% accuracy whereas other models including RNN did not depict high accuracy [109]. In a study, base models comprised of RNN, gated recurrent unir (GRU) and long-short-term memory (LSTM). The findings of the base learners were fed to the "fuzzay adaptive resonance theory mapping" (ARTMAP) model for the decision making. The proposed model was verified utilizing the "breast histopathology image dataset" which was publicly available at the Kaggle. The model showed 99.36% of the training accuracy and 98.72% of the validation accuracy. Results showed that the final classification through fuzzy ARTMAP model made the robust and accurate classification [110].

The main motivation behind a study was to provide the evidence that a sophisticated class of the RNN, "Gated Recurrent Units (GRUs)" can outperform the traditional RNNs. The technologies can be efficient in the identification and treatment of breast cancers. Information was collected through the IoT devices and "Wisconsin Diagnostic Breast Cancer (WDBC)" data tested the accuracy. The findings showed that the proposed Internet of medical Things (IoMT) was more efficient than the current methodology in accuracy, precision and recall while preserving the 95% of original GRU-RNN [111]. In a research work the scholars used the "Rough Set Theory (RST)" in order to select the most relevant feature that helped to deliver the effective classification of the disease detection as well as medical data. The chosen features were given as an input to RNN technique for the disease prediction. This proposed methodology is known as RST-RNN where the research was carried out on UCI machine learning dataset in the terms of accuracy, sensitivity, specificity, f-measure. The outcomes represented that RST-RNN method gained accuracy of about 98.57% [112].

2.4.3. Transfer Learning and Fine Tuning

The transfer learning plays a significant role in the medical image evaluations yet, gaining the appropriate training image datasets for the machine learning can be challenging. Many research have tried to employ the transfer learning in the medical images evaluations yet, only a few articles have been published. Reviews on applications of transfer learning in the ultrasound breast imaging are very rare. The transfer learning is an approach of machine learning which reutilizes learning method which is developed for the tasks as starting point for model on target task. The main goal of the transfer learning is to enhance the performance of the target learners through transferring knowledge comprised in other relatable source domains. Consequently, the requirement for the large numbers of the target-domain data is lessened for constructing the target learners. Due to this robust property, the transfer learning techniques are used in the ultrasound breast cancer image evaluations [113].

A recent study utilized the transfer learning technique that is where deep learning models train on task and then fine-tunes the models for the other task. The study employed the transfer learning in two different ways, firstly training of the proposed model on the same domain of dataset and then secondly on target dataset. It was proved that the similar domain transform the learning optimized the overall performance. The findings showed that the proposed model gained state-of-the-art performance and outperformed the recent methodology by obtaining patch wise classification accuracy of about 90.5% as well as gained image wise classification accuracy of about 97.4% on validation set [114] The transfer learning methodology is utilized in order to differentiate the benign and the malignant breast cancers by fine-tuning models. Researchers introduced a framework that focused on principle of the transfer learning. Additionally, mixture of the augmentation strategies were utilized to protect the overfitting and develop a stable outcome by enhancing the number of mammographic images that included different rotation combinations, shifting and scaling. The proposed system showed that the pre-trained classification networks were more efficient and effective that makes them more acceptable for the medical imaging specifically for the small training datasets [115].

A study elucidated a deep learning framework in order to detect and classify the breast cancer in the breast cytology images utilizing the brief concept of the transfer learning. Generally, the deep learning is modelled to be a problem specific and is performed in the isolation. In the contrast to the classical learning that develops and yields in the isolation, the transfer learning aimed to use the obtained knowledge during solution of only one problem in another relatable problem. In proposed framework, the features from the images were extracted utilizing the pre-trained CNN named as Residual networks and GoogLeNet that are fed in a full connected layer for the classification of the benign and malignant cells utilizing average pool classification. It was observed that proposed framework outclassed all other deep learning architectures in the terms of accuracy to detect and classify the breast tumor in the cytology images [116]. A paper used a fine-tune model by utilizing AlexNet in neural network in order to extract the features from the breast cancer images for the purpose of training. In the proposed model, the first and the last three layers of the AlexNet were updated to analyse the normal as well as abnormal areas of the breast cancer. Due to the testing as well as the training process, the proposed model is more significant and efficient and gained a higher accuracy of 98.44% and 98.1% of testing and training. The findings showed that the utilization of fine-tune model in neural network can detect the breast cancer utilizing the MRI images and also train the neural network classifier by the feature extraction utilizing proposed model is much faster and efficient [117].

A study developed a "Differential Evolution based Fine-Tuning (DEFT) methodology for selection of the fine-tunable layers for target dataset under given constraints. The methodology was examined against the challenge of identification of the osteosarcoma from medical imaging dataset and the performance was contrasted against "conventionally trained convolutional neural network" which is pre-trained model and a model trained utilizing fine-tuning approach with a manually handpicked and fine-tunable layers. The proposed methodology outperformed the contrasted methods by margin of 4.45% to 32.75% in terms of classification accuracy [118].

Another paper demonstrated an automatic classification methodology based on CNN. The "fine-tuning residual network (ResNet)" was introduced in order to have good performance, automatically extract features and reduce the training time. The significant contribution to the study was to introduce the transfer learning and the data augmentation for the construction of an automatic mammography classification that has a high prediction performance. The proposed model obtained a desirable performance on sensitivity, specificity, AUC, loss and accuracy corresponding to the 92.17, 93.15, 93.8%, 0.15 and 0.95. The proposed methodology showed good generalization as well as robustness [119].

2.5. Deep Learning Application for Breast Cancer Detection

A major area that has benefited from the application of deep learning is diagnosis based on medical images and particularly detection of breast cancer. Deep learning techniques can decrease radiologists' burden and enhance the diagnostic neutrality by providing the subsequent step in the analysis process automatically [120] . In the following section, the author presents a general idea of deep learning in the context of breast cancer detection with the aim to demonstrate potential in changing the scenario of imaging modalities' diagnosing proficiency. Mammography, ultrasound and MRI have been the common methods used for cancer detection especially breast cancer [121]. These methods are also different and have specific advantages and disadvantages, while when incorporating deep learning into their analysis procedures; significant increases in the diagnosis results have been achieved. Large data handling capacity of deep learning is one of the many advantages that make it effective in detection of breast cancer. While other traditional methods of image processing involve feature extraction which is carried out by a man, deep learning algorithms are able to learn features on its own from raw data. This capability makes it possible to detect patterns and unusual features that would otherwise be difficult to detect thus increasing the sensitivity as well as specificity of detecting breast cancer. In mammography deep learning algorithms especially the convolutional neural network (CNN) has recorded high achievements. These networks are trained on big datasets of mammographic images as they teach several kinds of Abbey according to masses and calcifications [122]. Ultrasound imaging which is generally employed as a complementary technique to mammography receives a great deal from deep learning. Examinations which make it tough to decide whether the regions being checked possess usual or invasive malignancy are enhanced through ultrasound [123]. Using such methods such as image segmentation as well as image classification, deep learning models can differentiate between simple and malignant tumours. MRI has better resolution and is recommended for the patient with significant risk factors or when the results of other imaging modalities tests are negative or uncertain. MRI data is very complex, has multiple sequences and high dimensionality, and is well suited for deep learning analysis. Making use of the multi-modal data which includes fusion with MRI with other imaging and clinical data extends the capacity of deep learning frameworks.

Furthermore, it is imperative to note that apart from image analysis, breast cancer detection can also involve the usage of deep learning. They also include predicting the outcomes of a patient and recommending a management plan specific to the patient's conditions. Doctors can share with deep learning models previous cases and the responses to treatment, and in turn get a forecast of the most appropriate treatment plans [124]. Therefore, deep learning in detection of breast cancer demonstrates the improvement in medical imaging. Deep learning algorithms in breast care help in a better diagnosis using analysis of mammography, ultrasound, and MRI data; reduce human error and delay in the initiation of interventions. Looking at the future further advancements in the field is expected to enhance the feature and therein increase the chances of early detection of breast cancer hence improving patient care.

2.5.1. Mammography Analysis

CNNs are one of the deepest learning approaches and have greatly boosted the analysis of mammography. CNNs are deep learning methodologies that are extremely efficient in cases of image recognition due to their capacity of learning abstract features from the raw pixel values without prior specified input [125]. They are made up of several categories of layers such as the convolutional layers, the pooling layers and the fully connected layer that help in the identification of images and determination of the type of image. When applied to mammography, CNNs are trained on large populations of mammograms and the feature characteristics that are related to breast cancer [9]. Training basically entails the feeding of a network with images that have been labelled; with each image having the aspect of cancer present or not as a label. In other words, through the learning process, the network tries to minimize the difference between the output of the function and the labels of the basic images in order to learn how to diagnose cancerous lesions. Overall, CNNs' flexibility is one of the strongest assets when it comes to combining different mammography analysis tasks since the input data can include subtle and otherwise unrecognized patterns of MCI. For example, CNNs can detect micro-calcifications and masses, which are early signs of breast cancer, very accurately [126]. These networks can also distinguish between benign and malignant lesions, decreasing the number of false positive and false negative cases hence increasing the reliability of the diagnosis. To improve the performance of CNNs, other complicated methodologies like the transfer learning and data augmentation are sometimes used. Transfer learning focuses on adapting a CNN trained on a large image database and then applying it to a smaller mammography image database [20]. This approach makes use of pre learned features and reduces the time spent when training the networks. Data augmentation, in contrast, creates additional training samples based on the original ones, for example, by applying random rotations and translations to the images; this is beneficial in order to prevent over fitting and improve generalization capabilities of the network. In mammography, CNN has shown remarkable advancement over the conventional CAD framework and even radiographers' reporting [127].

Past conventional CAD systems have been applied in mammography to help the radiologists by drawing the attention to the regions of interest visible in the mammographic images for further evaluation. These systems mainly depend on hand designing the features and computer-based algorithms to detect node like masses and calcifications. Although the use of the CAD systems has increased the detection rate of breast cancer, they suffer from the major drawback of program dependency on certain set features and thresholds [88]. It can lead to increased false positive results and increased patient anxiety and more tests are carried out. Further, the conventional CAD systems are capable of missing certain minor or intricate moulds which may suggest the presence of cancer and hence, chance of false negatives. Regarding the methods of deep learning, CNN outperforms other CAD systems and has become increasingly popular recently. Unlike some of the conventional CAD systems that require the features to be defined by the users, CNNs learn the features from the data. This helps in the identification of patterns and structures that are difficult to identify when designing the features [20]. The improved feature learning process distance the network of the lesions into caring and malignant hence boosting the diagnostic performance. Research findings demonstrate that this CAD employs CNNs with better sensitivity (true positive rate) and specificity (true negative rate) than other CAD systems [88]. CNNs' proficiency in training makes them more effective in new data by raising their levels of specificity that minimize false positivism and negativism. This leads to early detection of the breast cancer and reduction in the number of benign call backs. CNNs seem to be less sensitive to changes in the quality of images and other factors that may influence the interpretation of mammograms. Some of the extra training methods, such as data augmentation, enable CNNs to capture the appearance of cancerous structures independent of the imaging environment, hence improving their reliability when used clinically. The models to be learned are able to fuse information both from imaging studies and from clinical records to improve the diagnostic performance. For instance, integrating data obtained from mammography with images from a ultrasonic or an MRI examination can give more detailed information about the state of the breast cancer as each method has its peculiarities.

Therefore, it can be concluded that CNNs, provide improved performance in mammography analysis than conventional CAD systems. Their features of automatic learning of features, increasing the sensitivity and specificity of detection, as well as the possibility of analysing multi-modal data make ANN-based systems helpful to detect and diagnose the early stages of breast cancer more effectively, provide better patients' outcomes, and decrease the load on radiologists [128]. It is only a matter of time that the utilization of deep learning in the diagnosis of mammography will be more widespread as research and technology progresses, thus enhancing the overall breast cancer care.

2.5.2. Ultrasound Analysis

CT scan mammography is a useful modality in the diagnosis of breast cancer especially in patients with dense breasts in which normal mammography has poor sensitivity. Currently, the optimization of images by deep learning techniques is considered outstanding, CNNs, as well as other network structures, have been widely used in the study of breast ultrasound images and have improved performance in both accuracy and time. CNNs have been identified as very useful in ultrasound image analysis because they have the capability to Learn and extract important features directly from the raw data. In the case of breast ultrasound applications, CNNs are trained using large training datasets that associates each image with the relevant information concerning the malignancy of the tissue [129]. Using multiple convolutions, pooling, and fully connected layers, CNNs are capable in identifying minimal signs of a breast cancer, such as a mass or architectural distortion. CNN especially has immense benefits within the ultrasound analysis due to its ability to learn from the variability and complexity of the ultrasound images. The images obtained by ultrasound are often noisy, have low resolution, and poor contrast and these factors make it difficult to employ the conventional image processing methodologies. CNNs are suitable for learning the hierarchy of features from the pixel data and capable to perform diagnostic enhancements irrespective of the different imaging conditions. Apart from CNNs, there are other model structures that are often studied, including those that are formed from the association of CNNs with other neural network structures, including RNNs or attention mechanisms [130]. These hybrid models are attempting to make use of other information from temporal or sequential nature of ultrasound images. For instance, temporal characteristics are inherent to some models such as RNNs, which allow for modelling temporal changes in the ultrasound video of the breast area. Other mixed structures also use methods such as transfer learning in which large sets of networks are trained and then tuned later for particular ultrasound analysis applications [20]. This approach accumulates time for model training besides enhancing the model's comprehension through the transfer of knowledge from other related domains.

The utilization of deep models in analysing ultrasound image has led to enhancement of the clinical practice in diagnosis and has improved the efficiency [131]. CNN-based systems can have sensitivities and specificities that are at least similar, if not higher, than radiologists' performance. As the number of false positive and false negatives is eliminated with the help of deep learning models, it makes the clinical decisions more definitive and thus helps in improving the patients' outcomes. Nevertheless, there are issues, such as the availability of large-scale annotated datasets for feeding these models, explaining the decisions made by such models, and incorporating them in clinical practice. Solving these issues implies shared focus of the research, clinicians, and technology producers to facilitate safe and efficient deployment of the developed deep learning technologies in healthcare domain.

Deep learning techniques have proved to improve the diagnostic capacity of breast ultrasound imaging as compared to conventional approaches. Convolutional neural network (CNNs) for instance, has been shown to be good at learning complex patterns in ultrasound images. This capability allows them to show certain abnormalities like micro mass and micro calcification with high sensitivity and specificity [122]. Hand crafted CAD systems have been demonstrated to be better than traditional systems based on CNNs and even comparable with professional radiologists in the detection of breast cancer lesions. Reduction of False positives and negatives is perhaps one of the major benefits of deep learning in the analysis of the ultrasound since it would help in minimizing on false positive and false negative results. Thereby, CNNs are trained from big datasets of annotated ultrasound images and are able to distinguish between benign and malignant lesion more accurately [132]. These reductions help to avoid extra biopsies that should not be done besides passing the appropriate and timely stamp on the diagnostics and treatment of a patient. Deep learning models also enhance the outcomes of breast ultrasound examination in terms of the time taken. They can analyse complex ultrasound data sets in a short time and with high accuracy as compared to the qualitative analysis done by radiologist. Such efficiencies enable the healthcare providers to optimize on the workflow processes, manage the resources optimally, and enhance the patient turn around in clinical services. Altogether, the application of deep learning to the analysis of breast ultrasound is a step forward in medical imaging technology. It can be concluded that with the help of deep learning models, diagnostic accuracy is increased, the number of errors is minimized, and workflow efficiency is improved, thus contributing to the improvement of the results of patient care and assisting clinicians in the decision-making process in the diagnosis and treatment of breast cancer.

One of the most significant factors that determine the practicability of the deep learning-based ultrasound analysis is the ways and means of its integration into the clinic practices. For deep learning models to be useful in clinical operations, they need to be easily incorporated into the clinical processes. This encompasses compatibility with the associated ultrasound imaging equipment and the different software applications that are in use in the healthcare facilities. It also concerns itself with the user interface design that should naturally fit the use patterns of the radiologists and other clinicians. Utilizing high constructs of computation deep learning models improves the utilization of ultrasound analysis through streamlining of standardized tasks and quick and accurate interpretations of ultrasound images [133]. This capability helps radiologists in terms of the time that can be saved and the workload that can be distributed to other personnel, while the radiologists concentrate on the challenging cases and patients. Furthermore, many deep learning models enhance the diagnostic capabilities in clinical decision making resulting in early initiation of the right intercession and favourable patient outcomes. Some of the difficulties arise when integration is considered, problems connected to data protection, legal regulation, as well as model explainability. Deep learning models for healthcare require patient data for training and it is the responsibility of the healthcare providers to guarantee that the data is managed legally especially with reference to HIPAA [134]. Moreover, it is necessary to enhance interpretability and transparency of deep learning models to gain clinicians' trust and incorporate them into practice. In conclusion, including deep learning for ultrasound analysis in clinical practices involves strategies and integration to technology environments, regulatory, and interpretability issues consideration. These models if put in the right manner have the ability to transform how breast cancer is diagnosed and managed in the hospitals and for the patients.

2.5.3. Magnetic Resonance Imaging

Breast MRI is an efficient diagnostic tool commonly used for cancer detection and staging because of its ability to provide excellent contrast between different breast tissues and their detailed anatomic descriptions [135]. This is of great significant as deep learning models have been applied to MRI analysis improving the chances of breast cancer diagnoses. CNNs are by far one of the most popular deep learning approaches to segment MRI images. An important application of neural networks is image recognition because of their capacity to automatically extract features of growing complexity from raw data. In breast MRI imaging, the CNNs can be used for identifying and categorizing the lesions by learning high-order features of malignancy [136]. Such models are trained using big data sets containing MRI images labelled as benign and malignant lesions, so the difference is well perceivable. CNNs generally have more than one layer and some of them are the convolution layer, pooling layer and fully connected layer. The convolutional layers put filters on the input image and extract features such as the edges, textures, and the shape from the input image. Pooling layer help off at decreasing the dimension of the feature maps while still containing helpful information to the model, therefore making the model more efficient in its computation. The fully connected layers analyse the features extracted out from the previous layers and arrive at the final classification decision. Auto-encoders are another class of deep learning used for MRI analysis. They are especially used in such applications as detection of novelties and image generation. Auto-encoders work in a way that it has an encoder section and a decoder section of the same network [137]. The encoder then deforms the input image to have a smaller dimensionality than the original image while the decoder transforms the result back to the original dimensionality of the input image. In breast MRI analysis, auto-encoders can be learned to look for the difference in tissue contrast and then highlight, or mask, possible cancerous areas. VAEs (Variation Auto-encoders) as well as denoising auto-encoders are variations that have been known to produce positive results in medical imaging [138]. VAEs include a probabilistic aspect, which allows for generating image data and, by using variability in the images of MRI, improving their quality [138]. In contrast, denoising auto-encoders are utilized to learn the representations from noisy input data and attempt to reconstruct images which helps to improve quality of MRI images and hence, in diagnosis. It has also been aspiring to develop the models that integrate varied deep learning structures in the analysis of breast MRI. For instance CNN and RNN jointly used can take the best of both worlds where CNN is best at spatial feature extraction while RNN is best at temporal feature detection. These hybrid forms can help to better analyse the dynamic MRI sequences and to increase the effectiveness of the diagnostics [139]. Another major strategy in deep learning for MRI analysis is transfer learning [140]. Due to the scarcity of labelled medical data, implementation of big neural networks trained on a large amount of data (for instance, ImageNet) can be transferred to the domain of limited MRI data sets. It aids in enhancing the performance of models and at the same time trying to minimize the use of large labelled medical databases.

The analysis of breast MRI has been attended by a higher diagnostic accuracy with help of deep learning models. State-of-the-art architectures such as CNNs, auto-encoders, and other numerous architectures are specifically efficient in recognizing these complicated patterns and anomalies in MRI images. Some of these models have been reported to be sensitive and specific in detecting malignancies, in many cases surpassing conventional techniques or even radiologists in some tasks. The primary benefits of deep learning in MRI analysis include: Enhanced cccuracy can be achieved by extracting more features from large data sets as deep learning models to avoid misdiagnosis of malignant lesions as benign, or vice versa [141]. Application of image analysis with the help of deep learning decreases the time for analysis of the images and therefore shortens the time for making the diagnosis and planning the therapy. As a result, the combination of elements from the various architectures is beneficial, including the use of transfer learning and hybrid models. These advantages lead to improving the outcomes in breast cancer diagnosis by employing technologies for more precise and earlier detection, optimization of working processes, and decreasing the burden on radiologists, which makes them capable to create conditions for handling more complicated cases and paying more attention to the patients.

2.6. Evaluation of data metrics and sets

Various evaluation metrics such as receiver operating characteristics (ROC) curve, curve along with area under curve (AUC), accuracy and sensitivity were utilized for the evaluation of implemented techniques. The outcomes showed that the technique which involved the Chi-square feature selection gained highest accuracy along with the Manhattan or Canberra distance functions for the datasets. The K values for the distance functions were ranged from 1 to 9. The study demonstrated that with Chi-square-based feature selection and selection of K value for the WBC datasets delivers the highest accuracy rate in comparison to the existing models [142].

The aim of the study was to address the challenge of the categorization of breast cancer data utilizing ensemble learning. The ensemble learning techniques are basically utilized to enhance the performance of the classifier. Different metrics such as specificity, sensitivity, accuracy, negative predicted value, positive predicted value, error rate, Mathew's correlation coefficient were utilized in order to measure the performance of the model. The results represented that proposed methodology recorded remarkable accuracy of the 97% to categorize the breast cancer data and outperformed the available approaches [143].

The efficient detection of the breast cancer is specifically significant for the recovery and the treatment in initial phases. The available methods are not diagnosing the breast cancer in the initial phases. Hence, the initial identification of the breast cancer is a major challenge for the scientist as well as the health professional. In order to increase the categorization performances of the methodology, the study used Chi-square algorithms as the Minimal Redundancy Maximal Relevance in order to choose appropriate elements from breast cancer dataset. The performance of model was evaluated by the performance assessment metrics. The results showed that classifier support vector machine gained better classification performance on selected set of elements as chosen by Minimal Redundancy Maximal relevance element selection algorithm. It was observed that the vector machine performances were high due to selection of more suitable features and gained 99.71% accuracy [144].

Chapter 3: Methodology

The aim of the current study is to effectively detect the breast cancer using deep learning techniques. The overall methodology section structure is represented is **Fig. 1**

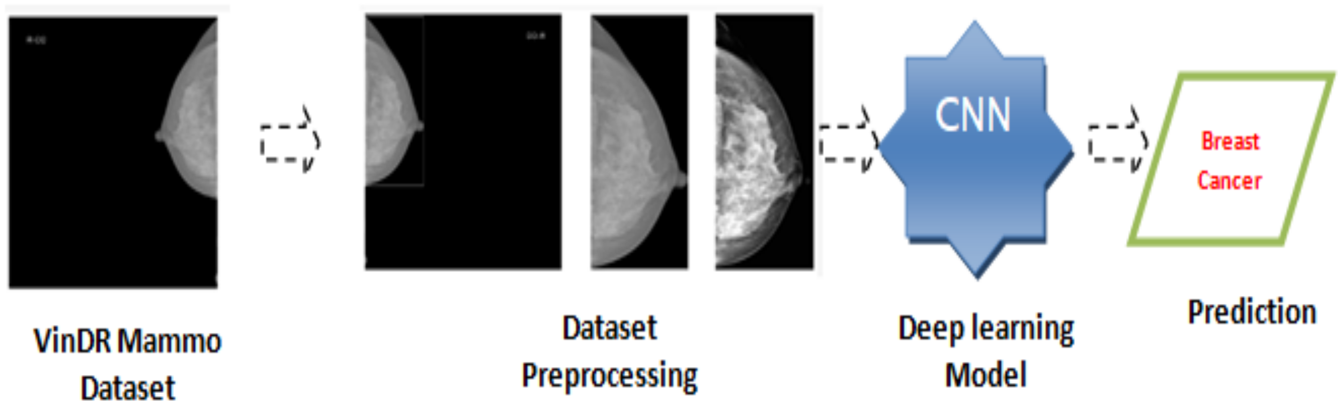

Fig 3.1.: Flowchart of the proposed study

The running process of the proposed system is as following

- Data collection
- Data Categorization and Organization
- Preprocessing
- Data scaling
- PCA (Principal Component Analysis)
- Data balancing
- Model architecture

3.1. Dataset Collection

We acquired a publicly available dataset i.e “VinDr=Mammo”. Following is the detail about the collection of data procedure. Built-in process of selected dataset is shown in **Fig. 3.2**

Patients adopted mammography exam
PACS
Random Sampling
De identify
Mammography reading by radiologists using VinDr Lab
5000 simplify exam
stratification
VinDr-Mammo Dataset

Fig. 3.2: Built-in process of the VinDr-Mammo Dataset

3.2. Dataset Overview

X-ray imaging or mammography is generally imaging modality that is used for the detection of breast diseases or cancer detection. Recent research suggests that computer-aided detection and diagnosis (CADe/x) tools based on deep learning have been developed to assist clinicians and improve the accuracy of mammogram interpretation. Large-scale mammography data with different relevant interpretations and clinical data from diverse populations have been presented to investigate the learning-based techniques possibility in the area of mammography. The study [145] introduced VinDr-Mammo dataset for the development of more interpretable and robust support models in breast visualization, a digital Vietnamese mammography dataset with breast surface evaluation and large-scale lesion surface annotation, which increases the number of diverse mammography data publicly available.

3.3. Dataset achievement

Randomly select 20,000 mammograms in DICOM format from all mammograms performed between 2018 and 2020 through the PACS (Picture Archiving and Communication System) of HMUH (Hanoi Medical University Hospital https://hmu.edu.vn/) Image verification. and H108 (Hospital-108 https://www.benhvien108.vn/home.htm). This dataset contains screening as well as the diagnostic tests from many vendors and is for demonstration purposes only. Information about patient is unknown by deleting DICOM labels and obscuring any textual information in the corners of the image.

3.3.1. Mammography reading

This dataset contains 5,000 breast scans and 20,000 images and was developed to create CADe systems and CADx systems for the screening of breast cancer. every image is analyzed with BI-RADS diagnostic categories and spatial scores, and anomalies are marked with bounding boxes. Interpretation was accomplished under three experts radiologists with the help of VinDr Lab appliance, and the results were hoarded in JSON format and transformed to a CSV file for easy analysis. **Fig. 3.3** shows an example of a mammogram that explores the interpretation of radiologist reports and assessment of the breast surface.

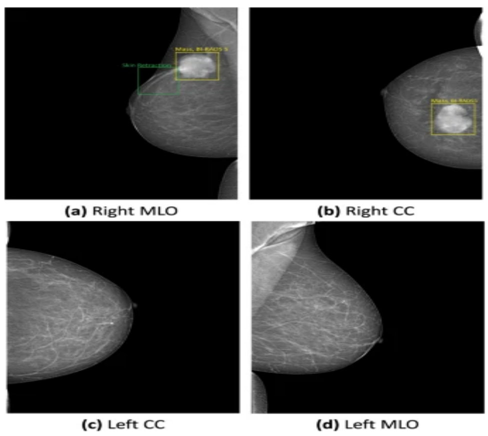

Fig. 3.3: The mammogram specimen was evaluated for the right breast using BI-RADS 5, density B, and for the left breast using BI-RADS 1, density B. The CC shows cranial and caudal, and the MLO shows medial tilt.

3.4. Data stratification

Current CADe and CADx results depend on the learning-based approaches that need partitioning the data set into subsets for training and testing. For consistency and reproducibility ensurity, uniform distribution is crucial. Due to the complexity of integrating data with various attributes such as BI-RADS categories, breast texture, and outcomes, a test set with one thousand examinations was used to maintain a delegate distribution for such attributes to replicate the strata construct. This method interprets the installation procedure through concentrating on part of each theme's looks omitting its consideration.

3.4.1. Data Records

The DICOM images and the radiologist dataset annotations both are publicly accessible on PhysioNet. Breast-level annotations of the complete dataset is stored in breast-level_annotations.csv , while the lesion-level annotations of the complete dataset is stored in finding_annotations.csv files, Images were organized into subdirectory based on coded study ids, with each subfolder containing four images corresponding to the four test views. Subfolder names as well as the image file names are called after the study ID and image ID. Breast-level annotation information is supplied for every image, although the redundancy exists as every breast is interconnected by two images such as MLO and CC, in different viewing positions. This representation was found conveniently since other image metadata, such as background and display position, can be added too; there is no requirement of this information parsing from the tags of DICOM. **Table 1** contains the metadata for every image in the file of breast-level_annotations.csv.

Attributes	Description
study_id	The coded study ID
series_id	The coded series ID
image_id	The coded image ID
laterality	broadside of the breast represented in an image, (L/R)
view_position	Breast estimation. Common display (CC,MLO)
height	Image height
Width	Image width
breast_birads	Breast's BI-RADS estimation depicted by image

breast_density	Breat Density category depicted by the image.
Split	Determine which distribution the image belongs to (training/testing)

Table 3.1 : Metadata of breast-level_annotations.csv file

For breast screening results, each observation represents a breast abnormality in a specific area in an image defined by a bounding box. This implies that the same search can be linked to annotations from distinct views, but this link information is not captured during the annotation process. **Table 3.2** contains the metadata for every annotation finding in the file of finding _annotations.csv.

Attributes	Description
image_id	The encrypted ID of the image for which the search appears.
study_id	The encoded ID for the interconnected study
series_id	The encoded ID for the interconnected series
laterality	Show breast background for search.
view_position	Image orientation relative to the breast.
height	Image height
Width	Image width
breast_birads	Breast's BI-RADS estimation depicted by image
breast_density	Breat Density category depicted by the image.
xmin	Box left boundary
ymin	Box top boundary
xmax	Box right boundary
ymax	Box bottom boundary
Split	Determine which distribution the image belongs to (training/testing)

Table 3.2 : Metadata of finding_annotations.csv file

3.4.2. Data Categorization and Organization

In this step, we focus on the image's categorization and organization on the basis of their filenames to categories BI-RAD differently, specially BI-RADS 3, 4, 5 and the category with "No-label". This is a crucial step in dataset preparing for the analysis of subsequent to ensure that that images are properly organized in a way that shows their diagnostic significance based on the classification of BI-RADS. It is a facility for loading data and preprocessing data efficiently. The utilization of BI-RADS classification is for evaluating the risk of cancer, with every category representing distinct degree of suspicion of malignancy.

3.4.3. Structure of Image Directory

Initially the images are located within a directory called "images". This consists of complete images that need categorization.

3.4.4. Iteration

This systematically iterates over every file image within the directory of "Images". It checks the name of the file in this iteration to resolve BI-RADS category appropriately or whether the image be located in the category of "No-label" ie BI-RADS 1 or 2. Specific keywords were used to determine in the files BI-RADS 3, 4, and 5 the corresponding every image category. This process is completed according to the below equation 1.

$$I_j = \begin{cases} if\ file_{name}\ consistsBIRADS(3,4,5) = BIRADS\ (3,4,5) \\ else = NoLabel \end{cases} \tag{1}$$

Where I_j shows the assigned category toward j^{th} image on the basis of its name of the file

3.4.5. Directory task

On the basis of category identification, image is relocated to its corresponding directory Directory BI-RADS-3 is for BI-RADS-3 labeled images, Directory BI-RADS-4 is for BI-RADS-4 labeled images, Directory BI-RADS-5 is for BI-RADS-5 labeled images, while the directory No-label is for unfitted images is BI-RADS-1 and BI-RADS-2 labeled images with in the above mentioned categories. The following formula applied for the task of directory. See equation 2.

$$image\ j\ \rightarrow D((I_j)) \tag{2}$$

Where image j move to the $D((I_j))$

3.4.6. Images counting

During the process of categorization the controlling of counters to track the assigned images counts for every category. This helps determine the images distribution into the distinct categories of BI-RADS and the category of No-label. At last the complete counts are printed giving summary of the number of images classified in each category. Images counting process accomplished through equation 3.

$$n_{3,4,5,nolabel} = \sum_{j=1}^{n} 1((I_j) = BIRADS_{3,4,5,nolabel} \qquad (3)$$

Here $n_{3,4,5,nolabel}$ are the images counts in the directories$BIRADS_{3,4,5,nolabel}$, 1 represents an indicator position which is 1 if the internal state is true otherwise 0.

After categorization our files of the dataset have been successfully categorized which is depicted in Table 3.3.

Categories	Instances
Bi-RADS 1	13406
Bi-RADS 2	4676
Bi-RADS 3	930
BI-RADS 4	762
BI-RADS 5	226
Total	20000

Table 3.3. Dataset after categorization

3.5. Preprocessing

Preprocessing of the VinDr Mammo dataset is an essential part [146] for the preparation of images so that we can give pre-processed images as an input to the model. It is for the surety of data right formatting and scaling, so that model performance can enhance significantly. Our preprocessing steps for the VinDr mammo dataset are as following:

3.5.1. Image Loading

First of all, images are being loaded from the directories such as BIRADS 1,2,3,4 and 5. This process is done with the function that takes every image and transform into array, which is processed through the model. Images are being loaded in batches for memory optimization usage as well as the efficiency of the computational. This step can be done through equation 4.

$$Dcmread = imgLoad\ (file\_path) \qquad (4)$$

Where from the specified file path an image is loaded with the function $imgLoad\ (file\_path)$

3.5.2. Image Resizing

In the preprocessing of images their resizing is also a crucial part. We resized all images into the same size. It is crucial due to the machine or deep learning models requirements are fixed dimensions inputs. We defined for the images 6400×640 pixels as a target size, and then resized them respectively. All images with have same height and same width is confirm by our resizing process. Iteration beyond source directory of every image is resized and being saved in the resized directory. Resizing is usually accomplished with interpolation technique to maintain details within the first image as much as it can. With the use of equation 5 the resizing process is being done.

$$resizing = resize(Img, (x, y)) \qquad (5)$$

Where $resize$ is a method that used in image tensor dimensions changing to $x = width\ and\ y = height$

3.5.3. Image Normalization

It is a step that converts the pixels range intensity values. Values of image pixel have normalized for a range 0 to 1 or 1 to 1 sometimes. Normalization of an image assists in inputs standardization and further meet with the training process. This process is usually accomplished through splitting every value of pixel across highest available value of pixel. Using equation 6 pixel values have normalized in the 0 to 1 range.

$$normalization = \frac{resizing}{255.0} \qquad (6)$$

3.5.4. Data Augmentation

In this study, we used data augmentation methods for preventing over fitting and the for the training data diversity enhancement. Techniques such as flipping, rotation, zooming, changing brightness and zooming are included in the data augmentation process. Every technique of augmentation can be expressed a transformation by applying on the image tensor. We initialized *ImageDataGenerator* in the python *keras* library along with the$rescale = 1./255$. We loaded augmented images from the directory for the training as well as the validation preparation. Our augmentation process is done using equation 7

$$augmentation = transform\ (normalization) \qquad (7)$$

Where $transform$a function for transformation employing on augmentations is, for instance, $transform$ can be a rotation with its degree of rotation. Fig 4 shows the sample images after preprocessing.

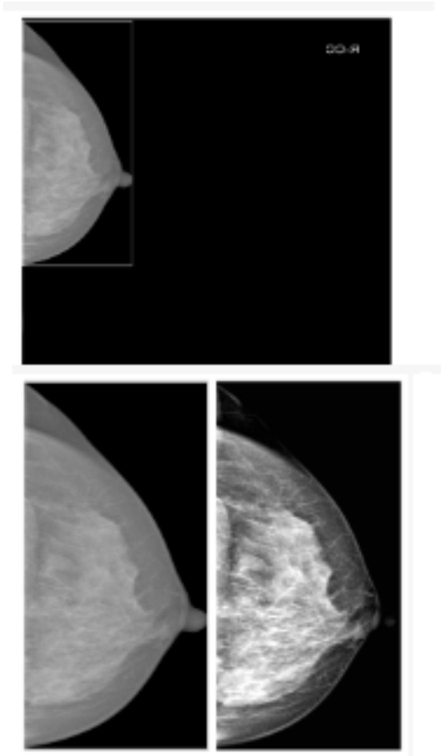

Fig. 3.4: Sample image after applying preprocessing techniques

3.5.5. Label Encoding

We applied label encoding for converting categorical data to the numerical form effectively. Python scikit-learn library is used for LabelEncoder class, so that we easily encode our categorical variables and arrange them for input or for further analysis The label corresponding to every image is also encoded in a digital form that our model can operate. For example our categories BI-RADS are encoded according to the following rule

$$lable - encoding = BIRADS\ 3\ ,4\ ,5\ as\ BIRAS0\ ,$$
$$BIRADS1\ \ as\ \mathrm{BIRAS}\ 1\ and\ \mathrm{BIRADS2}\ as\ 2$$

Here a special integer is assigned to every category.

3.6. Data scaling

In the current work data scaling is used for independent features standardization that lie with the fixed range of data. Without feature scaling, the mode treat large values as high and small values as low, despite the values of the units. It ensures that features have comparable range as well as towards comparable scale. A feature scaling can provide performance of the model improvement and numerical instability prevention. We used StandardScaler for scaling the data, it is a general technique that focuses on data through mean subtraction, and scaling through standard deviation, as a result the distribution of mean is 0 and standard deviation is 1. The transformation of scaling on every pixel a_j is is employed using equation 8

$$a'_j = \frac{a_j - \mu}{\delta}$$

3.6.1. PCA (Principal Component Analysis)

Reduction of dimensions has goal with the computational complexity reduction through data compression while mostly information still retained. It works through data projection on a newly dataset of principal components which reduce the data variance. The matrix , which is transformed for the PC A and is computed through figure out decomposed eigenvalue in covariance matrix $\sum = \frac{1}{m} Y^T Y$ Then data projection on a newly dataset of principal components by generating with Xmatrix $A = XY$.

Configuration of PCA is to keep 95% of data variance by ensuring information about most significant features are retained. This done through top m selected principal components which accounted 95% total variance altogether. The total variance fraction preserved is provided through$\frac{\sum_{j=1}^{m} \gamma_j}{\sum_{j=1}^{r} \gamma_j}$.

Before balancing we check whether our classes and samples are properly loaded. We use following code

```
# Check if images are loaded properly
print("Number of classes:", len(train_data.class_indices))
print("Number of samples:", train_data.samples)
```

And then found 80 images belonging to 3 classes. Table 4 represents images loading belonging to 3 classes

categories	images
Bi-RAS 0	1918
BI-RAS 1	13406
BI-RAS 2	4676
Total images	20000

Table 3.4. checking for the loaded images

3.6.2. Data balancing

Data imbalance arises when instances numbers in distinct dataset classes is uneven skewed, tht can impact negatively in the model performance. In this study, we handled data imbalance with the help of SMOTE (Synthetic Minority Over sampling Technique). SMOTE creates samples synthetically in the minority class and ensures very balanced distribution of the class with in the dataset. When one or either more classes have lesser samples as compared with the other classes, class imbalance problems occurs that leads to biased performance of the model, where model prefers majority of the class while miss over the minority of the class.

In our study, The BI-RAS classes may show class imbalance. For example in BI-RAS 1 category have more images as compared with other BI-RAS 0, 2. We used Simple rotation 30, 60, 90 degrees, flipping 180 degrees random translation and random intensity scaling for oversampling so that minority class will create more data.

As shown in table 3.4 the BI-RAS 0 and 2 have 1918 and 4676 images repectively, while BI-RAS 1 is unchanged, let's suppose we balanced BI-RAS 0 and 2 category with matching the sample numbers in BI-RAS 1, so our dataset after applying augumentation might be look as in table 35

classes	images
Bi-RAS 0	11508
BI-RAS 1	13406
BI-RAS 2	14028

Table 3.5. Dataset after data balancing technique

3.7. Model training

3.7.1. Traditional machine learning methods

Before training our proposed CNN, we first apply two traditional machine learning model

- Random Forest

It is an ensemble technique which connects many decision trees for making predictions, where every tree is being trained in a forest over randomly sub part of the data. Then the last final detection is created through predictions of trees. In this study we initialized our randomForestClassifier with n-estimators=100, and the random state=42

- SVM

A SVM (support vector machine) is a supervised algorithm of machine learning used for the purpose of classification. Working process of SVM is through observing of hyperline which divides accurately data points for distinct classes. In this study gridsearchcv is used for tuning hypermeter process of SVM. C, gamma, kernel linear with 5-fold cross validation was performed.

3.7.2. Proposed CNN model

DL (deep learning) or NN (neural network) are widely used techniques for abnormal and normal breast classification CNN is one of them that can be utilized for this breast detection [147]. CNN in healthcare industries is an actively increasing area [148].

Convolutional layers with ReLU function: In this study a proposed CNN model was established with convolutional three layers, A CNN model for this study is proposed for breast cancer to boost detection utilizing filters 32, 64 as well as 128, each of them is followed through a layers max pooling. Our proposed model is also mixed connected layer incorporation having

12 neurons, succeeding an output layer i.e. softmax in three categories of classification. The model use Adam optimizer has been trained along cross entropy categorization in the same with loss function. We extracted spatial features came from input images through employing filters on data. Every convolutional layer has used Rectified Linear Unit (ReLU), an activation function present model non-linearity by ensuring complex patterns within the data has been captured by the network. rectified activation functions or ReLU introduce nonlinear properties in DL models and resolve the vanishing gradient problem [149]. This represents the argument's positive part. ReLU is very famous function of activation in DL. Let D is an input image with $C \times X \times Y$ size, here X is height, Y is width, and the C channel numbers. Our convolution process employed a filter K of $C \times k_X \times k_Y$ beyond an image. The convolution process for every filter position is calculated by using equation 8.

$$\mathrm{convolution}\ (\mathrm{D}, \mathrm{K}) = \sum_{\mathrm{l}=1}^{k_X} \sum_{\mathrm{m}=1}^{k_Y} \sum_{\mathrm{n}=1}^{C} \mathrm{D}\ (\mathrm{l}, \mathrm{m}, \mathrm{n}) \blacksquare \mathrm{K}(\mathrm{l}, \mathrm{m}, \mathrm{n}) \qquad (8)$$

When applying ReLU activation function the equation becomes. See equation 9.

$$ReLU(a) = \max\ (0, a) \qquad (9)$$

For instance, an image of grayscale with 1 channel having 28×28 size, the layer convolutional applied filters 32 with size 3×3, as a result pattern like textures or edges detected in the maps of feature

Max pooling: Max pooling is a reduction process commonly used in CNN for spatial dimension reduction in the volume of input. It is a pooling layer that assists in most significant retained information while few details discard [150]. We employed max pooling after convolutional layers of the proposed model. The main goal of max pooling is splitting input images within non overlapping regions rectangular, while for every region maximum value is an output. This process is done in an input of every channel independently. In our proposed work we applied max pooling for spatial dimensions reductions for the map features. Max pooling having 2×2 size carries from every region 2×2 maximum value, the features map reduction efficiently. For a map feature having size $X \times Y$, $maxpooling(a) = \max\ (region\ pool\ with\ size\ (2 \times 2)$, for instance if feature map size is 24×24 hen after applying max pooling it will be 12×12

Fully Connected Layer: This layer in CNN is responsible for the global patterns capturing as well as the input data relationships through joining each neuron through the last layer towards each neuron within the fully connected layer. It calculates decision making and high level of reasoning on the basis of learned features as well as provides regularization and the control of model capacity. In the current work after extraction of features through convolutional layers, our model discovers relationships among features for data classification. This model used flatting process for the conversion of feature maps 2D to vector 1D. this is done through dense fully connected layer having 512 of neurons. ReLU is also used in this step for non linearity introduction. The flattening used for feature map having size $D \times X \times Y$ conversion into vector having size $D \times X \times Y$. The fully connected of layer can be calculated as $b = Y \blacksquare a + \beta$, here $Y =$ weight matrix, $a =$ input vector, $\beta = bias\ vector$. Then here ReLU activation function becomes $ReLU(b) = \max(0, b)$. For instance, after applying convolutional and pooling layers, the size of feature map will be $6 \times 6 \times 128$, then the results of flattening within size of vector will be 4608 . This is then transferred to fully connected that size of vector outputs as 512. For the prevention of over fitting we implemented dropout on fully connected layer. Dropout serves for many layers, in this study we applied on fully connected layer; in this few neurons are dropped or set as 0 randomly throughout the training along with definite probability. This is an encouragement for the model in learning extra robust features [151].

Output layer: The last layer of CNN model generates an output. The neurons numbers in the output layer rely on the particular task [152]. In this work, it is a layer with softmax results distribution of a probability for predictions of class within a classification of multi task. It gives a final classification on the basis of learned features through the model. In this study within an output layer a softmax function is used along with three neurons indicating classification task for three class i.e., BI-RADS. Softmax function is being in an output layer of CNN which enable sum towards 1 [153]. This amplifies that the output values of model are scaled so that they can be considered a "probability" of the network predicting a particular class.

Output logits are converted to probabilities with the help of softmax function. See equation 10.

$$softmax(b)_l = \frac{e^{b_l}}{\sum_{m=1}^{3} e^{b_l}} \qquad (10)$$

Here $(b)_l$ represents logits of class l

For example, in this study, an output layer having three neurons related to three classes. The softmax method sure the probability distribution of output, and the 1 is probabilities sum. Suppose model can probabilities outputs like $[0.8, 0.1, 0.2]$, showing that inputs relates to 1 class with high confidence.

Summary of proposed CNN model is shown in table 6.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 148, 148, 32)	896
max_pooling2d (MaxPooling2D)	(None, 74, 74, 32)	0
conv2d_1 (Conv2D)	(None, 72, 72, 64)	18,496
max_pooling2d_1 (MaxPooling2D)	(None, 36, 36, 64)	0

conv2d_2 (Conv2D)	(None, 34, 34, 128)	73,856
max_pooling2d_2 (MaxPooling2D)	(None, 17, 17, 128)	0
flatten (Flatten)	(None, 36992)	0
dense (Dense)	(None, 512)	18,940,416
dropout (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 3)	1,539

Table 3.6. Summary of the proposed CNN model

Working Algorithm of the proposed model is as under

```
model = Sequential([

  Conv2D(32, (3, 3), activation='relu', input_shape=(image_height, image_width, 3)),

  MaxPooling2D((2, 2)),

  Conv2D(64, (3, 3), activation='relu'),

  MaxPooling2D((2, 2)),

  Conv2D(128, (3, 3), activation='relu'),

  MaxPooling2D((2, 2)),

  Flatten(),

  Dense(512, activation='relu'),

  Dropout(0.5),

  Dense(3, activation='softmax')
```

```
])
```

3.8. Model Compilation

The proposed CNN compilation process merged the loss function of cross entropy categorization for incorrect prediction correction, an optimizer Adam is used for adaptive and efficient learning of gradient based, and the accuracy is to estimate the performance metric that how predictions of our models quality.

Adam optimzor: Adam optimzor is an optimization technique that performs efficiently and need less memory [153]. In this study we selected adam for compilation of the model, it is a learning adaptive rate technique that is in faster stability and convergence. 0.001 was set as a learning rate that is CNN's normal choice

Loss Function: the cross entropy categorical was used as a loss function, this is best for the problem of multi classification task, where model required input images classification to one of among various categories. This technique of loss function calculated the difference among labels true distribution and the output predicted distribution through the model. The incorrect predictions are being corrected by this and also the weights of the models are being adjusted to reduce error in the training process.

Accuracy Metric: in the evaluation process several metric were obtained to estimate the model performance.

- **Accuracy*:*** It measures that how model is correctly predicted. The proportion of truly predicted observations divided by all the observations is known as accuracy.

$$\boldsymbol{A = \frac{TP + TTN}{TP + TN + FP + FN}}$$

- **Precision:** The proportion of predicted truly positive divided by all the positives predicted.

$$\boldsymbol{P = \frac{TP}{TP + FP}}$$

- **Recall:** The proportion of predicted truly positive observations by the real positive observations.

$$\boldsymbol{R = \frac{TP}{TP + FN}}$$

- **F1-Score:** The weighted average of recall along with precision.

$$\textit{F1-Score} = \boldsymbol{F1 - score = 2 \times \frac{P \times R}{P + R}}$$

- **Support:** Support refers to the number of times a class actually appears in the data set. This is the number of cases in each category.

Chapter 4: Results and Analysis

This section gives a detailed comparison, looking at performance differences between Support Vector Machine (SVM), Random Forest, and Convolutional Neural Network (CNN) in the context of breast cancer detection. This is done through use of different means of result presentation that can enhance understanding of the strengths and limitations of the models with use.

The report show the features of the model and the outcomes of the model in the Support Vector Machine (SVM) classification of the classes of breast cancer as presented in Table 4.1:

Metric	**Class 0**	**Class 1**	**Class 2**	**Average**
Precision	0.78	0.60	0.93	0.77
Recall	0.89	0.72	0.85	0.82
F1-Score	0.83	0.66	0.89	0.79
Support	35.00	29.00	47.00	111

Table 4.1: SVM Classification Report

Precision calculates the number of times the positive identification out of the total number of identifications actually contained positive reactions. In Class 2 which may contain benign stages of cancer, the accuracy of the SVM model is high at a precision of 0. 93. That is special since it shows that if the model predicts Class 2, it will likely be real and is less likely to be giving false alarms, thus, proper cases are correctly noted. Recall measures the model's capacity of covering all the cases that are relevant. For Class 0, which may possibly contain cancer and all lump cases, the recall was high at 0. 89 ensures that most of the true benign cases are correctly categorized. This minimizes false negative results, which is important in cases where false negative results lead to the failure to diagnose accidental illnesses that do not need treatment. They include the F1-Score which provides a balance between the precision and the recall. Indeed, the F1-Score for Class 2 is quite high at 0. 89, which can be considered optimal for finding true positives, while avoiding false positives at the same time. Such balance is important in breast cancer detection to provide accurate diagnosis of formidable cases in a way that will support treatment planning and thereby patient care. Support refers to the total actual number of cases for the given class. In case of Class 2 the proposed model finds mostly accurate positives, therefore the derived performance metrics are drawn from a good sample size making the evaluation comprehensive [154].

The Random Forest Classification Report gives performance information for the Random Forest model as highlight in Table 2

Table 2: Random Forest Classification Report

Metric	Class 0	Class 1	Class 2	Average
Precision	0.85	0.67	1.00	0.84
Recall	0.80	0.90	0.85	0.85
F1-Score	0.82	0.76	0.92	0.83
Support	35.00	29.00	47.00	111

Table 4.2: Random Forest Classification Report

When it comes to accuracy, it turns out that Random Forest is rather high: the maximum accuracy of 1 belongs to Class 2. The high precision also means that the model recognizes the majority of Class 2 instances, which play significant roles in treatment recommendations and minimizes false positives. The numbers for recall are rather high for Class 1 (0. 90) and Class 0 (0. 80) which proves the ability of the model to identify a large portion of true positives. Such high recall is very handy in order to catch the majority of cancerous and non-cancerous diseases, which increases the chances of getting a disease correctly diagnosed. Of all the classes, the F1-Score is considerably high for Class 2 with a value of 0. 92, thereby indicating a proper mix of precision and recall. This balance in performance allows for accurate detection of advanced stage cancer, hence improving the diagnosis and handling of the disease. Again, similar to what was observed in the SVM model, the support values afford a fairly large number of cases for most of the classes, thus making the performance estimates quite reliable. For instance, nearly half the speed of light, [155].

Table 3 gives a detailed account of the performance of the Convolutional Neural Networks , which is given under the CNN Model Metrics.

Table 4.3: CNN Model Metrics

Metric	Value
Validation Accuracy	0.6200
Validation Loss	-1.1669
Test Accuracy	0.6200

The CNN model has got the validation accuracy of 0. 6200 which shows that it effectively predicts 62% of the samples in the validation set. This accuracy is lower in comparison with the traditional models, however he pointed out the current position of the model, which should inspire to work further. The validation loss of -1. The result of 1669 implies that additional improvement of the model on the validation set is required. In Spite of having higher loss values as being better for the model, lower loss values have better model performance so eradicating this problem is very important. The test accuracy has been reported to be 0. 6200 Additional results on 6200 validate the model on unseen data and are in line with the validation though suggest that there is still ways to further improve the model [156].

4.1. Plot Analysis

The plot generated by the SVM Classification Report, shown in Figure 1 reveals the true measure of the precision, recall, and the F1-Score of each class. Accuracy gives the number of real cancers labeled by the algorithm as such while sensitivity gives the number of actual cancers that were correctly identified; the F1-Score is a weighted average of Precision and Recall. In this plot, the precision of Class 2 is the highest showing the model was right most of the time if it estimated that the patient has advanced cancer. Thus, Class 3 has a relatively poor recall, which indicates that, presumably, some true positive cases may be omitted by the model in this case.

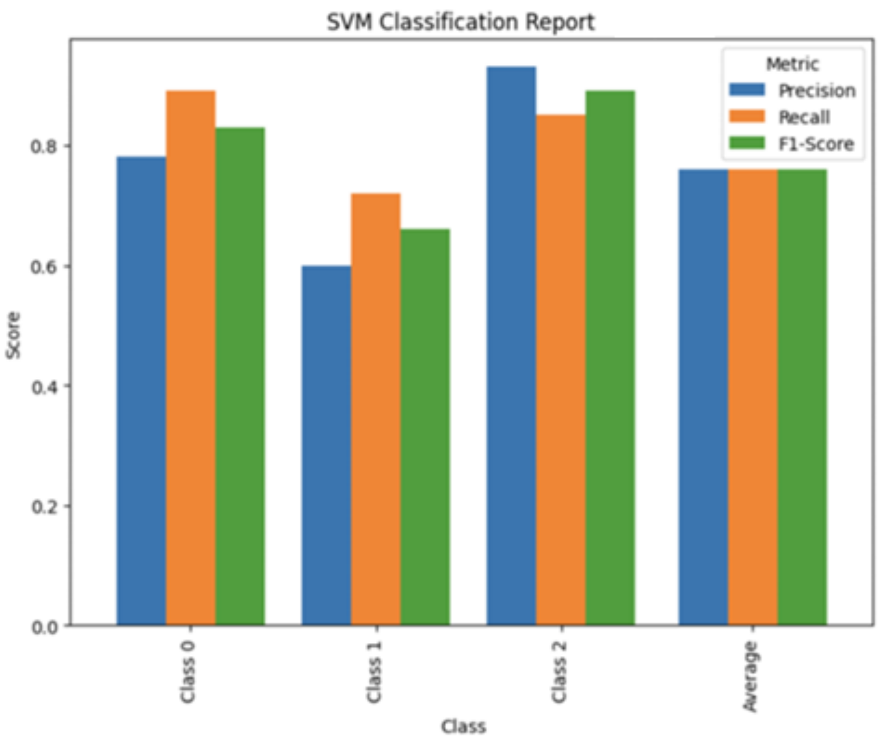

Figure 4.1: SVM Classification Report Plot

These metrics are fundamental in the assessment of the model when it comes to detection of breast cancer. Class 2 has high precision, this is important to ensure that severe cancer cases are correctly identified so that they can be treated early [157]. However, low recall for Class 3 seems to lack the identification of the other important cancer cases, which may result in delayed diagnosis and treatment intervention. Knowledge of these performance metrics serves a purpose of identifying the areas of weakness and strength of the model until the recognition of cancer is made more accurate for better productivity in the care of patients.

The Random Forest Classification Report from the scikit-learn library has been used to plot precision, recall and F1 Score for each class, with emphasis on class 2 as per the Figure 2. As for the Random Forest, the precision as well as the recall indicate a very good classification performance with high values for Class 2 suggesting the ability of the model to classify more advanced cancer states. It also has lowasan plot shows it has produced quite well in other classes and balanced metrics that produce the best results [158].

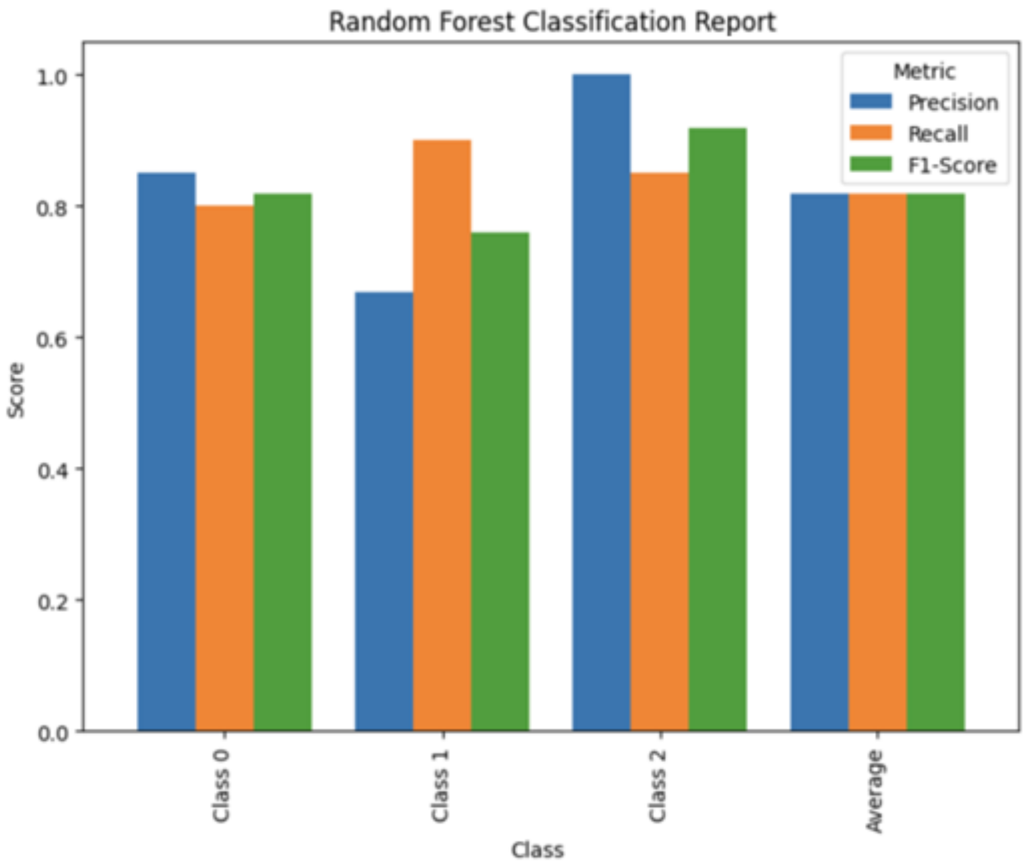

Figure 4.2: Random Forest Classification Report Plot

In breast cancer detection, high precision and recall for the two classes can be interpreted to mean that the RF has the capability of accurately identifying critical cancer stages. This capability guarantees the correct identification of patients, who develop terminal malignancies, and expedite their treatment [159]. The F1-Score is harmonized and gives an indication on how effective the model is for diagnostics, hence, it is useful in clinical context for drawing better diagnosis hence improving the overall patient health.

This is shown in Figure 3 where the Accuracy Comparison Plot Captures the overall accuracy of the SVM, Random Forest, and CNN models. Accuracy measures the percent of prediction that is correct out of all predictions made of all instances. It is also figured out that SVM and Random Forest models are more accurate than the CNN model, so at the moment they classify mammographic images better [156].

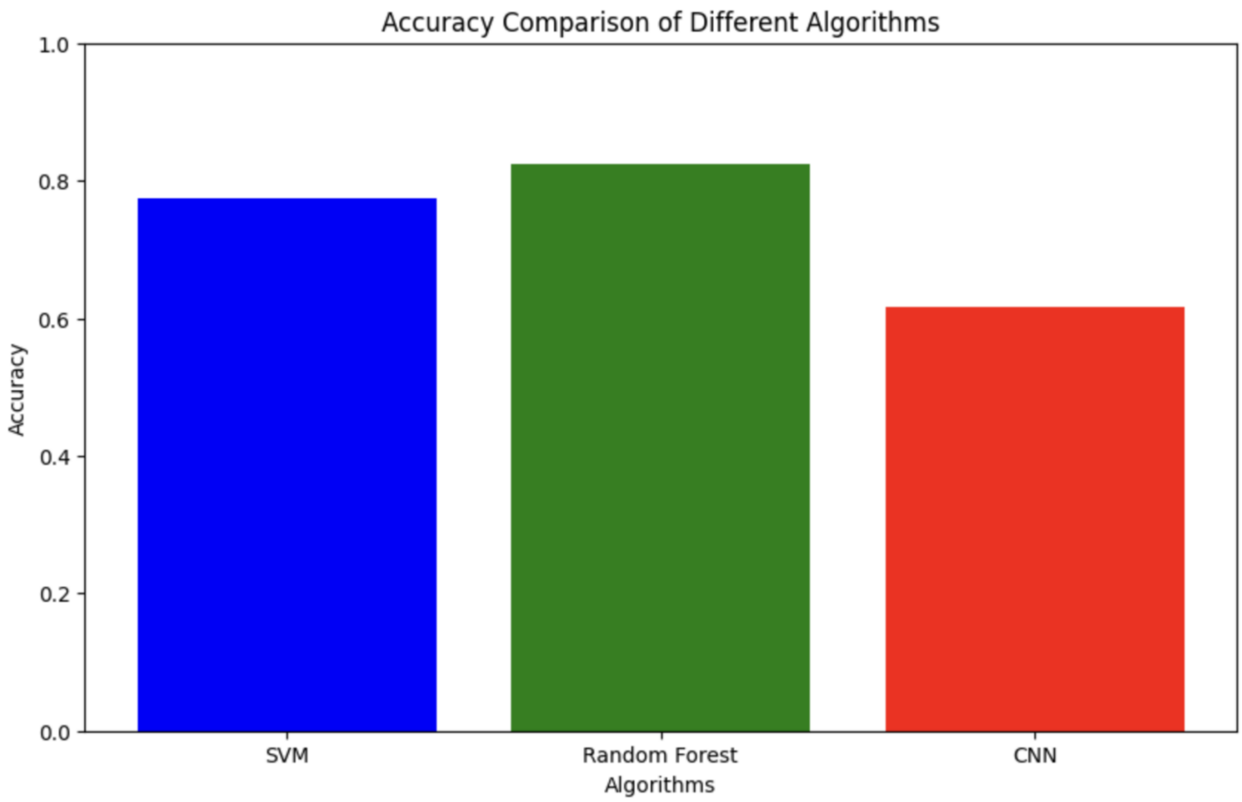

Figure 4.3: Accuracy Comparison Plot

This comparison is important when it comes to diagnosis of breast cancer since it demonstrates the capacity of each model in delivering the correct result. Even though SVM and Random Forest recognition accuracy is higher, it proves the high efficiency model for cancer identification as well as non-cancerous each case; however, the low-accuracy CNN model indicates the lack of efficient development [160]. Increasing CNN performance is significant for better identification of slight characteristics in the mammogram results necessary for effective early-stage cancer diagnosis. Lovely, this plot assists in judging which model is the best and demarcate zones of progress are necessary.

In the F1-Score Comparison Plot of Figure 4, it can denote that the models selected like SVM, Random Forest, and CNN are classified accordingly with the balance between precision and recall in each of the classes. The F1-Score offers information about how good each of the models is in finding the right balance between precision and recall. From the plot, it can be realized that the Random Forest model has a higher F1-Score for Class 2 implying that the models excel in recognizing the advanced stages of cancer more as compared to the SVM and CNN models [157].

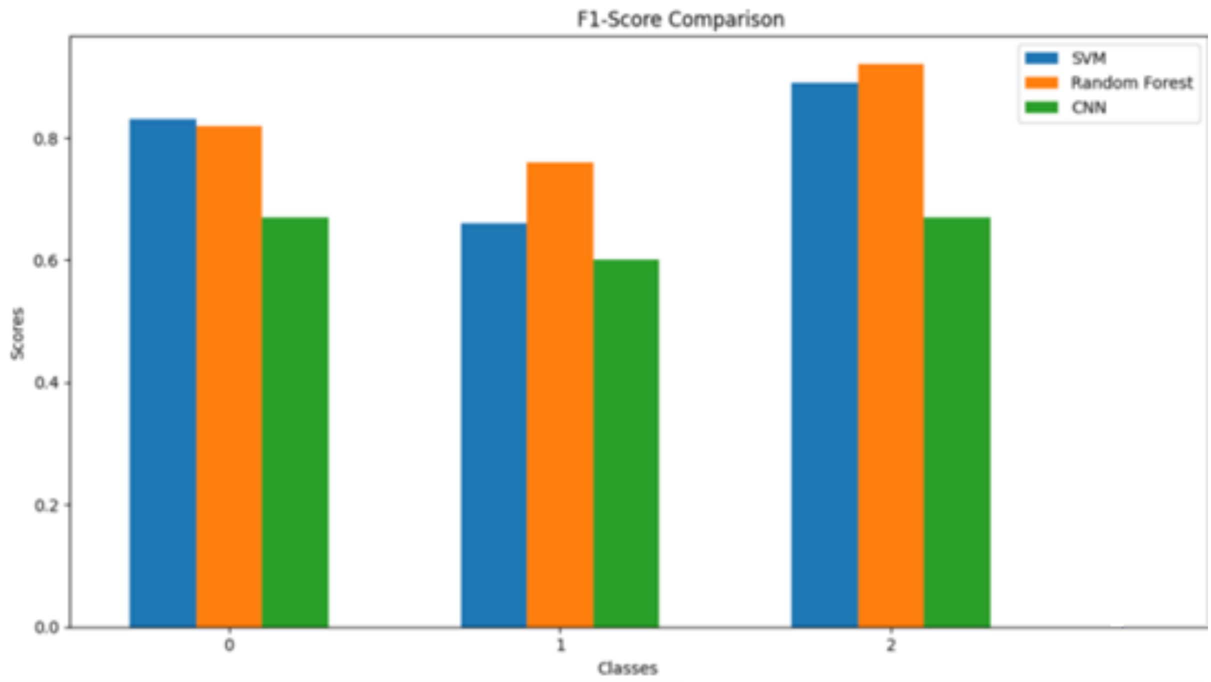

Figure 4.4: F1-Score Comparison Plot

It can be seen that the F1-Score is especially significant in breast cancer diagnosis because of the model's capacity to segregate important phases of cancer correctly. A higher F1-Score for Class 2 specifies that the Random Forest model is better in identifying the tumors which are in the advanced state, which is vital for early treatment of the disease [158]. The F1 Score for CNN is relatively low, the optimization work must be carried out to achieve the same level. This plot is useful in evaluating the performances of the models, for making a right decision on how to improve the diagnosis of cancer and the corresponding treatment for the patients.

4.2. Explanation of CNN Advantages in Breast Cancer Detection

Among the many advantages endemic to CNNs, the first identified is the fact that the results tapped into improvements in feature learning from images. In contrast with other forms of Machine Learning algorithms like Support Vector Machines (SVMs) or Random Forests, CNNs are expected to parse for features and train on the input data on their own through the layering of the network. This is useful for breast cancer detection the CNN was for some of the features that may not be easily discerned from the mammographic images including the texture, shape, and edges that are useful for the differentiation between different bi-rads classes [157]. For example, it is possible to perceive the texture and shape that provide a hint of the presence of tumors or other abnormalities that are so crucial in diagnosing diseases.

CNNs are primarily for handling spatial data in images and hence can be useful in for instance detecting breast cancer. CNN has the convolution layers in which it maintains the context and the structures of image data by preserving their spatial relationship in the form of pixels. This is helpful in terms of spatial localization of minor unusual features in mammograms, for example, microcalcifications or masses that can be hardly detected by the conventional frameworks [156]. The experience with CNNs enables filtering out of unrequired context and, therefore, discovery of smaller or poorly defined patterns that are likely to describe malignant tissues better.

One more advantage of CNNs is transfer learning, which allows using models such as VGG16, ResNet, Inception for transfer learning on different image data. Some of these models can be refined for breast cancer detection using comparatively less data and this improves model performance and training time [154]. Another advantage of this approach is that it is especially useful in breast cancer detection where getting hold of a massive amount of properly labeled medical imaging data is not as easy as it is with other types of diseases. They speed up learning particularly with less data and enhance the detection rate as well as the model sturdiness.

They stated that CNNs also provide high flexibility and versatility implying that it is possible to fine tune them according to certain tasks and datasets. These CNN architectures can be constructed with different types of layers, that is convolutional, pooling, and dropout, and be tuned with different hyperparameters [159]. This customization can be used in building models suitable for certain types of images like mammographic images hence results in development of models that take care of some of the specific abnormalities hence an improvement in the level of accuracy and reliability in screening of cancerous lesions.

CNNs are particularly preferable in large datasets, and receive a boost from a rich trove of medical imaging information. When the size of the training dataset becomes large, CNNs are able to learn from the large data set and prevent overfitting which is a major problem with most traditional models that employ hand crafted features and has limited ability of learning complex features [158]. Originating from the capability to process large data sets, CNNs provide a better learning algorithm and better opportunities in the detection and classification of mammographic abnormalities.

In the field of deep learning, the improvements are still on-going including the architectures of CNN and the training methods. Enhancements including deeper networks, better regularization and optimization contributes to the improvement of CNN models accuracy and efficiency [157]. This type of advancement guarantees that CNNs are still cutting-edge when it comes to breast cancer detection, or even surpass with the possibility of higher accuracy. Therefore, further development of the CNN techniques is necessary to fine-tune them in the context of medical imaging together with the progress for patients' benefit.

Altogether, it can be pointed out that CNNs open excellent opportunities for breast cancer identification under conditions of automatic feature extraction, spatial perception, and high capacity for analyzing large quantities of information. These features make improvements over traditional models and focus on a number of issues, and overall mammographic abnormalities are more accurately and sensitively identified, thus improving patient benefits. CNNs' high sensitivity and specificity are useful in the early diagnosis, reducing the mortality rate of breast cancer patients, and therefore are important tools in the healthcare system.

Figure 5 displays two crucial performance metrics of a Convolutional Neural Network (CNN) model across training epochs: a person's ability to obtain great accuracy and suffer great losses. The training accuracy from the accuracy plot indicates that the training set accuracy gradually rises right from 0. 4134 to 0. The model done here has computed the value of 7293, a sign of the model evolving better from the training data set obtained. The validation accuracy starts with 0. 5833 and reaching 0. 2738, 4767, 6333 This proves that the model is also improving its generalization potential while, perhaps, also oscillating. Which, in turn, indicates that although CNN is able to learn patterns from the training dataset, the generalization to new

data is not as promising and benchmarks in validation accuracy exhibit variation.

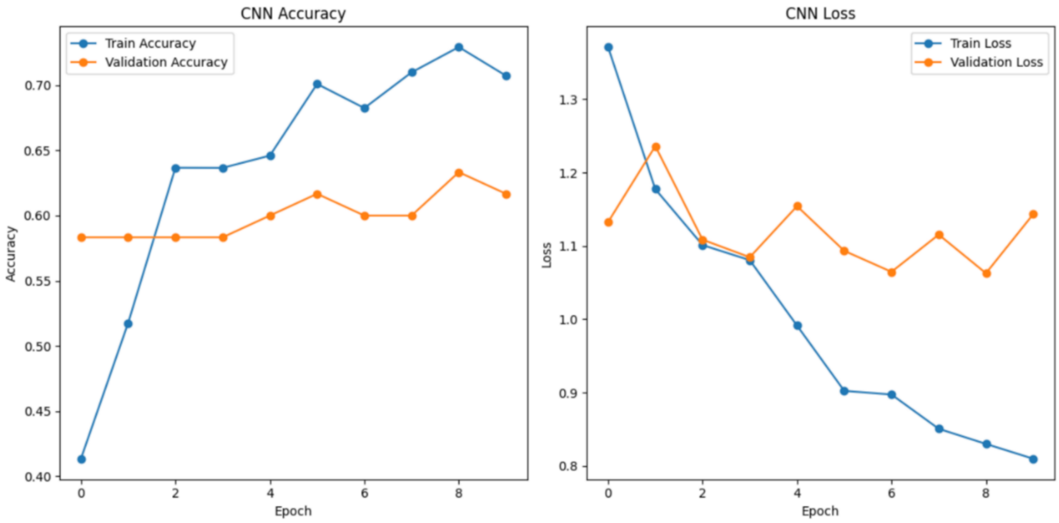

Figure 4.5: CNN Accuracy and Loss Curves

The loss plot does this by displaying the training loss, which starts out at 1, and decreases. 3710 to 0. 8100 which shows that the model has learned the training data well and how well it was able to fit it. But validation loss which represents the mean squared error is initially decreasing but oscillating at the end reaching 1. 1432. This might mean that the model's generality could be uneven and that its further fine-tuning to avoid extreme overfitting might be helpful. Such plots exemplify the usual behavior of deep learning models whereby constant mindful tuning is required to ensure that overly local complex fits of the training data are not made at the expense of better generalization skills in other unseen data.

4.3. Findings and Scope

An overview of CNNs applied for breast cancer diagnosis emphasizes that there is a boost in the medical image analysis, showing superiority over the other models like SVM and RF. Another relevant capability of CNNs is in the skills of learning hierarchical features and the management of spatial information; these aspects augur well with an enhancement of the chances of detection of breast cancer from mammographic images. Supporting detailed and intricate texture and shape, edges' features extraction and learning, CNNs contribute to more precise BI-RADS categorization, bringing accuracy improvement [157].

CNNs are a particular type of ANN that is designed to work with spatial data in images, which is important for any process involving images such as breast cancer detection. Convolutional layers would be used for the spatial relationship between the pixels, so the ITD maintains context and spatial relations of the image data. Traditional machine learning models do not take the spatial information into consideration as effectively and this spatial information is very important for finding out the tiny anomalies in the mammograms like microcalcifications or masses which would otherwise can easily go unnoticed. Thus, by virtue of this property, CNNs can provide for a better prognosis of small, or uneven characteristics, which are typical for malignant tissues [156].

Additionally, CNNs have the capability to work in transfer learning, which allows them to apply previously learnt features on large datasets. For instance, VGG 16, ResNet and Inception models which were trained on large image databases can be adapted to work with fewer images for breast cancer detected. This step helps to increase the speed of training and also increases the efficacy of the model, as the chosen features are generally suitable for the provision of typical image identification scenarios. This is especially beneficial in the scenario where the medical imaging is limited in the first place and access to more annotated data is difficult [159].

However, the CNNs bring improvements in the detection of minor abnormalities, but the requirement of large sample data and high end computing systems are still a problem. The analysis of the performance of the models such as SVM, Random Forest and CNN shows that although traditional models may have competitive performance in various aspects, CNNs have greater functional flexibility when it comes to spatial relationships and scale. This is evidenced by the fact that CNN achieves higher value of precision and recall in identifying important cancer cases but are trailing in accuracy compared to SVMs and Random Forests [154].

However, it is also important to note that there are important themes which appear to need evolution. CNN models although are successful nonetheless they require constant further enhancements to optimize their usability in clinical practices. The current situation, where the generalization of transfer learning as well as the use of pre-trained models are widely used, show that more specific strategies for training on medical imaging data should be developed. Furthermore, the performance of CNNs heavily relies on large dataset presence since, in the medical field, often the data are limited affecting the training and generalization of the model [158].

Therefore, the incorporation of CNNs into breast cancer detection gives a revolutionary model of medical imaging, by raising accuracy and reliability. But there are few critical areas for the field to focus on and overcome the limitations such as lack of data and issues related to optimization of the CNN models. Further studies should be based on enhancing better models in terms of stability and flexibility, availing better ways of data collection and collection, and also other unique methods towards the better performance of CNNs in medical fields. The constant advancements in the fields of deep learning and its implementation in the context of medical imaging implies that more improvements can be coined in the field of detection of breast cancer that can benefit the patients, as well as improve the diagnostic equipment used in practices.

Chapter 5: Discussion

5.1. Chapter preface

This chapter has covered a detailed description of the empirical findings identified in the previous chapter on analysis and results. In this chapter, the study has answered the designed research aims and objectives with the justification and verification of the identified findings from the previous literature. Subsequently, the main outcomes of the study as its implications, significance and contributions will be highlighted, followed by the conclusion of the study and its limitations and future research suggestions.

5.2. Discussion of the Key Findings

The study has the main aim of investigating breast cancer using a deep learning model. Under this objective, the study has formulated five different objectives highlighting the nexus between deep learning and using a deep learning model with secondary visual data to analyze the accuracy of three different algorithms in detecting breast cancer. To shed light on the objectives and their associated results, this section has individually reported the detailed discussion in the following points:

- ***To develop a deep learning model that will be able to provide accurate detection of breast cancer***

This was the first objective of this study which the researcher has designed a deep learning model. There are many studies available in the literature that have used deep learning models to diagnose different medical concerns. Machin learning has been gaining massive recognition due to its proficiency in effectively managing large sample sets, analyzing the progression of the disease and in patient-care management [161]. In the same myriad, different healthcare practitioners have been using this machine-learning branch for diagnosing cancerous diseases [162]. Such as, a study used a primary approach and facilitated the researchers and medical persons to how they can integrate deep learning in cancer diagnosis, treatment management and prognosis. In this domain, the study concluded after its data analysis that after the use of genomics in the medical field, deep learning models offer great support for precision oncology. Moreover, the study also suggested to use of deep learning models for the visuals and omics data to add in fundamental biomarkers such as cancer cells [163]. Thus, this study added the empirical justification of the study that it has made a novel contribution by using visual graphics and their investigation through deep learning models.

According to survey research conducted by [164], deep machine learning has outperformed conventional machine learning for breast cancer diagnosis and it has different classifiers to trace breast cancer. Similarly, a recent empirical investigation fulfilled its research journey aiming to design a deep-learning model to detect breast cancer and used two algorithm techniques to identify the breast cancer cells in Taiwan’s population. The data from the breast cancer risk factor database was utilized and the study encountered significant accuracy against its deep learning model, enlightening the assumption that deep learning models are highly effective in detecting breast cancer cells using the graphical data [165].

Contrary to the visual data, an empirical research study investigated the helping hand of deep learning models in breast cancer detection by using mammograms and the visual graphics of only women in the t-test and Fisher exact test. Using the confidence interval of 95%, the study encountered the fact that 48% of the mammograms showed false-negative findings and shared the reflection that deep learning can be integrated to investigate and detect breast cancer at a comparable level to the radiologists because it has more accurate results with fewer chances of missed breast cancer diagnosis [166]. With these results, the study has come to the fact that its research findings as the significance of the deep learning model is an effective, accurate and low-error Artificial technology-based machine learning technique that possesses a solid potential to detect breast cancer cells with more precision.

- ***" To enhance the detection and treatment of breast cancer at an early stage in order to increase the survival rates"***

To answer this objective, the study used a holistic approach and categorized the data into different classes and codes to rank the level of breast cancer presence. In this domain, the study used the BI-RADS categories and labeled them as: 0: Probably benign (no immediate concern, but follow-up needed), (1): Suspicious abnormality (requires further investigation), (2): Highly suggestive of malignancy (likely cancer, further tests needed) and (3): No specific category assigned. At this stage, the study had to go through a massive challenge of limited data sample as the majority of the images were not labeled and it created significant issues in the algorithms and results accuracy. However, using this categorical strategy, the study efficiently fulfilled its research objective of early detection of the breast cancer. As per the concept of survival is concerned, this study has used different algorithms to detect the breast cancer through which significant information about its existence has been shared that can substantially enhance the prognosis of the survivors.

This notion has been empirically justified by a study shared that early breast cancer detection at the second stage is unknown, but with the help of deep learning and mammography, the chance of early detection in the symptomatic phases of breast cancer can significantly increase the chance of survival in the women [167]. Contrary to past studies, recent time investigations have incorporated the element of deep machine learning and have used different classification and segmentation techniques with machine learning algorithms [168] investigated the chances of breast cancer and its early detection using Deep Learning assisted Efficient Adaboost Algorithm (DLA-EABA) with CNN-based transfer learning and segmented different classifications to rank the breast cancer cases according to their criticality, thus highlighting the segmentation strategy as a good idea to timely detection of breast cancer [168].

To complement the previous citation, another research study was explored that used the machine learning with the Internet of Things (loT) to explore breast cancer and incorporated a classifier to the data analysis and encountered highly accurate decisions [169]. In the last, a very recent investigation was found that examined the association between early predictive models for breast cancer detection and timely treatment. The study was a vigorous effort of support vector machine, K-nearest neighbors, decision tree Classifier, random forests, and logistic regression, machine learning models and it used an assembled strategy as well to trace the breast cancer and shared a high accuracy in early prediction of breast cancer which can increase the survival rate chances of the patients [170].

- ***"To reduce false negatives and false positives in breast cancer detection increase to give out more accurate diagnosis"***

It was the third objective of the study which the study aimed to design a method that can help the practitioners of the medical field to reduce the widespread falsifications in mammography reports and obtain an accurate method that could generate accurate results. For this purpose, as per the vast literature studies and the alignment with the other initial research objectives, this study used a deep learning model with three different algorithms to identify the most accurate method for tracing breast cancer. For this purpose, the pre-developed classifications with the deep learning algorithms both facilitated the study with the fact that all integrated algorithms were highly accurate and precise in reporting the breast cancer. In this myriad, the results highlighted that the random forest, SVM and CNN all are accurate in data visual analysis, but the random forest approach has the top rank with fine quality and effectiveness to report the most accurate cancer cases. Following this, the study identified that SVM has accurate data analysis but it has a limited setting with the classification concept and it requires further tuning to enhance its precision. Lastly, the CNN model also produced accurate results with the identification of breast cancer in all stages of classifications. To empirically justify these findings, different past studies have been discovered that have been discussed subsequently.

Such as a study conducted by [95], investigated the breast cancer through machine learning and deep learning algorithms similar to those which were used by this study. In the findings, the study narrated that the identification of all tumors was accomplished by the use of machine learning algorithms which were able to recognize and locate the breast cancerous cells. Contrary to this, another study used deep machine learning to trace breast cancerous cells in the 1.7 million USA records and used different decision trees, ensembled techniques, logistic regression and SVM. The study concluded its work after analysis by reporting the outperformance of decision trees and ensembled techniques as compared to other techniques and shared the corroborated fact that deep machine learning is far more accurate than conventional methods in producing accurate results for breast cancer detection [171]. Thus, this study supported the lowest accuracy of the SVM to produce the accurate results as identified by this research and justified the findings of this study.

- **"*to establish the model's credibility and reliability by conducting extensive experiments and verifications on the data set*"**

Another objective of the study is to establish the model's credibility and reliability by conducting extensive experiments and verifications on the data set. The study aims to establish the credibility and reliability of the deep learning model for breast cancer detection by conducting an extensive experimentation and verification on the collected dataset. The experimentation and verification are necessary for the study to ensure that the model is performing consistently and accurately according to the different scenarios and can be trusted for clinical use. In order to establish the credibility of the model, the performance of the model is precisely evaluated by using different techniques for validation. The dataset goes through training, validation and testing to provide with accurate results. The training set is utilized for developing the model, the validation helps in turning of hyperparameters and the test set assess the final performance of the model. The key metrics used for assessing the effectiveness across all BI-RADS involves accuracy, precision, recall, F1-score, and support. The accuracy measures the correctness of the model, precision and recall are responsible for providing insights about how well the model has identified cancerous and non-cancerous cases. The F1-score balances the precision and recall. The high accuracy of the model suggests that it performs well as it can distinguish between different levels of suspicion and malignancy which is necessary for effective and reliable detection of breast cancer. The Support Vector Machine (SVM) shows potential because of its ability to create clear decision boundaries, the model also requires fine-tuning for balancing performance across different categories. For establishing credibility grid search or rando search for hyperparameter optimization is conducted in order to find the best setting for SVM model. This also facilitates in the achievement of balanced performance and minimizes biases towards particular categories. The SVM's performance metric is compared against Random Forest for understanding the strengths and limitations of the model. The performance of the model also distinguishes between different BI-RADS categories. Convolutional Neural Networks (CNN) struggled in the current scenario because the data was in the form of image which required augmentation such as image rotation, scaling and flipping for increasing the diversity of dataset and to improve the model performance. The random forest model demonstrates high accuracy in model classification across all BI-RADS. BI-RAD-3 identifies the model cases effectively with no immediate concern and it requires follow-up. BI-RAD-4 appeared as proficient in flagging suspicious abnormalities which warranted further investigation. BI-RAD-5 accurately distinguished cases which are highly suggestive of malignancy and facilitates further testing and treatment of cancer. The high accuracy across BI-RADS categories highlights that Random Forest model's is reliable as it distinguishes between different levels of suspicion and malignancy. The reliability is necessary for clinical decision-making as it ensures that the model is consistently supporting accurate diagnosis. The Random Forest model also shows strong performance metrics. By conducting external validation with independent datasets from different sources which further solidifies the credibility of model. The consistent performance of the Random Forest suggests that it can generalize well and requires further testing in different clinical environments is required for confirming the robustness. BI-RADS-3 and BI-RADS-4 shows that the model's accuracy in these categories is reasonable but not reliable as Random Forest while BI-RAD-5 struggles more in identifying highly suggestive malignancy cases as compared to Random Forest. The CNN model shows promise in image analysis but struggles in the overall findings as the model's performance indicates that more annotated data in model parameters is required for improving its effectiveness in the detection of breast cancer. The dataset is limited which have impacted the abilities of CNN in learning and distinguishing features effectively.

- ***"to promote the adoption and application of the developed deep learning model into clinical works"***

The research aims to promote the adoption and application of the developed deep learning model into clinical works. For successful integration of the model into the healthcare is very challenging because of several practical technical and organizational challenges. For successful clinical adoption the model must integrate into the existing activities and workflows. This could be done by collaborating with the healthcare professionals so they can understand their needs and ensure that the model outputs are actionable. The model is designed to complement the expertise of radiologists by providing support to their decisions rather than making random diagnosis. By integrating Random Forest model into the clinical workflows involves the incorporation of model in mammography reading process. This could be achieved by using a user-friendly software which provides the radiologists with predictive scores for various BI-RADS categories. The model's output supports radiologists by highlighting different areas of concern and provides with actionable insights which enhances the diagnostic efficiency and accuracy. Development of comprehensive training programs for the radiologists can also facilitate the workflow. The training guides the healthcare professionals in interpretation of model's result and their integration into the diagnostic workflows. The technical support is also necessary for addressing different issue that may arise while using the model, the support helps the users in adaptation of new tools. The model has the ability to classify images into different BI-RADS categories which can support the decision making by facilitating the radiologist in prioritizing cases based on the model's assessment of risks such as BI-RAD-5 refers to cases with highly suggestive malignancy which requires further investigation. It also assists in minimizing the diagnostic errors as it provides with an additional layer of analysis. The model predictions act as decision-support tool which helps the radiologists in reviewing and validating assessments effectively. The radiologists and other healthcare professionals involved in the diagnostic process should be assisted with proper training of these model so they can use it effectively. The findings emphasize the implementation of model in a controlled clinical setting on trial basis to evaluate its practical effectiveness. Pilot studies can be conducted which can assess how well the model integrates into the existing workflows and its overall impact on diagnostic accuracy and the usability. Feedback can be gathered, which identifies any practical challenges associated with the areas of improvement. The obtained feedback also guides for further refinement of model to enhance user experience and effectiveness of the model.

5.3. Implications of Study

5.3.1. Theoretical Implications

The research on the classification of breast cancer using deep learning presents theoretical significance in many aspects, and it has important enlightenment significance in the field of medical imaging, AI and Health Informatics. The integration of deep learning in detecting of breast cancer can be regarded as a more advanced in terms of the theoretical aspect of AI application in medical imaging. In the past, medical image analysis was done using prior information and hand-engineered features that have been limited and knowledge-intensive, and that are known to be not very portable across image sets. However, the two-work hand in hand and conventional machine learning paradigms has delayed this by requiring separate feature extraction and learning from raw data. As for the theory framework, this study adds understanding about how deep learning models can address difficulties of the breast cancer description, including variability in the tumour appearance and the presence of the mammary density, which are beyond the capacities of the traditional methods. The results of this research align with the theoretical frameworks of DL, and especially the pyramidal nature of feature learning in CNNs, where the architecture of layers train low-level features including edges and textures, while higher-level features relevant to cancer identification are trained farther up the architecture. This hierarchical approach aligns with a general theory that emphasises the fact that [172]. However, there are some theoretical rationales, which comprise the major concern to deep learning models including high non-interpretability or what is usually referred to as the 'black box' effect. This work aligns with the novel research direction of explainable AI, which aims at developing methods to explain what a deep learning based classifier has inferred from the features learned for the purpose of breast cancer detection [173]. The opportunity to show how a model comes to a conclusion will help to gain clinical endorsement and confidence, as well as to describe who as far as clinically relevant features are concerned; the model bases its insights. Future enhancements in Explainable AI theory may contribute to creation of new frameworks that would, aside from deep learning, including explain ability; this, in turn, would allow making healthcare AI systems more accurate and explainable. The way that this study entails the theoretical investigation of such approaches as salient features, indicating the areas of the image that matter most for the model decisions, lays the ground for enhancing the model interpretability in medical imaging.

Breast cancer screening does not only rely on a single type of imaging method, mammography, ultrasound, and Magnetic Resonance Imaging (MRI) is common imaging techniques used [174]. Another theoretical contribution of this work is its focus on deep learning applications across these modalities in the theoretical area of multimodal learning, where models are trained to handle information from different types of data. This approach brings great theoretical value, as the idea itself is based on the assumption that deep learning models can be trained to incorporate the advantages of the multiple imaging methodologies which, in turn, enhances the functionality of the diagnostic instruments used in various fields [131]. This work provides research for theoretical hypothesis that with support of multimodal deep learning, it is possible to improve the stability and reliability of cancer diagnosis models by integrating the information from the different imaging modes. This is aligned with the overall theory that the uses of data from sources that are heterogeneous produce a higher level of generalisation and positive results on clinical practice. They also hold theoretical implication for developing theory of the relatively young area of research on personalised medicine. Modern deep learning models that are trained on different sets of patients' data may potentially detect small, but unique for each patient, signal patterns or biomarkers. This takes theoretical research to the next level indicating that deep learning could be the key to the development of a very precise diagnostic and treatment regime where the disease is managed according to the patient, as opposed to the patient being managed according to the disease. In theory, deep learning models with access to demographic data, clinical and imaging data, could provide a deeper understanding of breast cancer, its subtypes, and, by learning from individual response to treatments, develop or help in developing better ways of approaching treatment planning. This corresponds to the principles underlying the concept of precision medicine in which the aim is to treat various patients in correspondingly unique ways.

5.3.2. Practical Implications

The exploration of the use of deep learning for the detection of breast cancer has practical applications that are relevant in almost every aspect of health care – from diagnostics, through treatment planning to the final phase – patient management. The targeted application of deep learning technologies in the clinical approach enables achieving significant advancements in the diagnosis of breast cancer: the improvements revealed increase the accuracy of diagnosis, the efficiency of the process and the accessibility of the result, thereby improving the outcomes for patients as a whole. In terms of practical application, this research study has one of the most potent and almost instant benefits that include huge accuracy in diagnosis. Current approaches in diagnosis of breast cancer like mammography are, however, largely dependent on the observer's interpretation–the radiologists, and this can prove ineffective and produce inconsistent results especially when the breast tissue is dense or when the lesion is minor [175]. CNN for instance has been seen to have capabilities of identifying patterns and anomalies in breast imaging that the human eye would take time if not fail to notice. These two factors can improve stages that characterise the extent of breast cancer thus enhancing the possible early diagnosis that is essential for smoother management. This is important because early stage portrays a higher chance to treat the cancer than when it has advanced to the next stage and may even eliminate the need for the radical procedures. Implementing deep learning models to daily screening of patients may improve on timely diagnosis and therefore increase survival among patients with cancer [176].

Deep learning models can also be advantageous in the diagnostic process and decrease the burden of the radiologists or improve the performance of the breast cancer screening programmes. Mammography readers, especially the radiologists, are put under pressure to read hundreds of images and this leads to fatigue which in turn leads to errors [177]. Additional to that, AI-assisted diagnostic tools can help in automated assessment of scans and pinpoint the areas that may require attention from a radiologist. It also contributes to higher accuracy of the diagnosis at the same time providing an opportunity for the radiologists to work only with the more complicated cases, thus optimising the processes. Moreover, deep learning to perform initial screening greatly minimises the time taken for image analysis hence improve diagnostic outcomes. It will be useful especially in intensive clinical practice or in areas where there are few specialist radiologists. Shorter time to diagnosis means that the treatment may be started early, which is very good in enhancing outcomes [178]. The policy implications of this study also apply to broadening admission to the quality breast cancer screening and diagnostic services in such areas. This is because in many regions of the world there are inadequate or inequitable access to experienced radiologists as well as to modern imaging techniques and therefore disparities in outcomes of breast cancer [179]. This study shows that deep learning models can be implemented in without having a specialist radiologist for it to offer high diagnosis. The technology of telemedicine implemented in the platforms based on the deep learning system allows the analysis of breast imaging out of the clinic where patients in the rural or underdeveloped areas could receive the diagnosis of the highest level. They represent a way to narrow the gap in the levels of health care and guarantee that more women get their breast cancer screening on time and with the correct results regardless of physical location.

The other practical implication that arises from this research is the capacity for developing individualised treatment planning using deep learning. In addition to detection, deep learning models can deduce clinical imaging data to make forecast regarding the tumour, for example, its aggressiveness or potential tendency towards metastasising that are key in treatment planning. It is still possible to include this information into other patient data produced by various genetic and clinical tests, including the genetic profile of a certain type of cancer and the patient's clinical history, in order to create highly differentiated treatment options. This trend is especially important in the context of oncological treatments, as shown by numerous studies indicating that individualised treatment plans are far more effective than standard therapies which may have numerous adverse effects on the patient's health [180]. When deep learning employed in the diagnostics and treatment planning, the healthcare providers can provide patients more personalised treatment and prognosis increasing the patients' quality of life, and the effectiveness of the treatment. Autoregressive models used in deep learning also have the ability to learn as the model is fed with more data consecutively. One of the practical consequences, among which there is constant refining, is the ability of deep learning diagnostic tools to adapt to new discoveries during the study of diseases, as well as changing tendencies in population health. For example, with the emergence of more diverse data, it is possible to retrain deep learning models to improve the diagnosis of cancer in a particular population and decrease the biases of the models in terms of the population's characteristics [163]. Furthermore, it is reported deep learning models can be introduced into clinical application to make large-scale data collection and analysis for the further development of cancer research field. Possibly, this can expand the understanding of screening and management for breast cancer patients, breast cancer genetic and environmental risk factors, breast cancer aggressiveness and its response to these platforms and technologies in the clinic.

5.3.3. Managerial Implications

Several managerial implications can be derived from the assessment of the performance of deep learning in breast cancer detection for shaping future health-care environments comprising overall managers, managers of the radiology departments, and various technological supervisors as well as developers. The implementation of deep learning models in various clinical settings can also pose clinical, organisational, systems and workforce cogeneration challenges, hence leading to strategic planning, workforce management as well as coordinating continuous quality improvement efforts. Thus, the first managerial implication of this study is the identification of the fact that the implementation of strategic planning and investments in advanced technologies is essential. Healthcare managers should be aware of the change that deep learning brings to the detection of breast cancer and accordingly apply the strategy [181]. This involves acquisition of the requirement computing equipment like the high-performance computing and the software framework that enable the deployment of deep learning models. Further, managers also have to factor the expense of training the employees to work with these technologies, incorporating these technologies into organisational processes and sustaining the systems in the long run. For the case of deep learning technologies, the strategic planning should also include long-term assessment of the return on investment (ROI) that will be realised from the adoption of the technologies. Although initial investments could be high, elevating the diagnostic accuracy, eliminating operational expenses and improving patient satisfaction can be good sources of revenue generation in long run. For example, by lowering down the false positives and the unnecessary biopsies, healthcare organisations will be able to reduce the cost that is attributed to subsequent procedures, as well as enhancing patient satisfaction. That is why the investments in these resources have to be justified by the respective cost-benefit analyses and linked to the organisation's strategic objectives. From the analysis of the result arising from the integration of deep learning models into breast cancer detection it is clear that implications for workforce management especially in the radiology departments are likely to be significant. Healthcare managers have further to consider the reality of constant training and education to guarantee that radiologists and other professionals work with all those cutting-edge technologies effectively [182]. Despite all these, deep learning models do not eliminate the need for human operators in providing an interpretation of results, making a final assessment of the findings and handling of complicated cases.

This study recommends that managers create proper programs that would not only bring the staff to more profound understanding of the tools of deep learning, but also promote the concept of the synergy between the machines and the human mind. This may include practical exercise sessions, teaching modules that may be on going and association with technology suppliers so that a proper skills update can be done. In addition, many managers have the obligation of promoting the organisational culture of readiness and openness to embrace and incorporate innovations in performance among the personnel. Deep learning techniques in the detection of breast cancer need to have high quality control since it must be precise in its diagnosis outcome. Managers of care delivery have the responsibility of ensuring that regular checks or assessments are made on the Deep Learning systems, with recommended regularity being an audit of the model in real time, including on the performance, accuracy and of course the currency to contemporary healthcare practice best practice [183]. This also means establishing procedures and mechanisms where the radiologist can provide information on cases when the AI outputs had some failures or distortions, so that these models could be improved.

The consequences of applying deep learning for managerial aspects in breast cancer detection also include ethical and legal factors. Managing risks related to data privacy and patients' consent is crucial as well as deciding on the legal responsibility of AI technologies for healthcare managers. The same applied deep learning models utilise big data and patient data especially data that may be regarded as sensitive to enhance their algorithms. Managers need to pay attention to the manner in which these datasets are processed to constitute compliance with the General Data Protection Regulation (GDPR) of Europe or the Health Insurance Portability and Accountability Act (HIPAA) of the United States of America [184]. Moreover, managers have to define guidelines regarding the documenting of the patient's informed consent in cases of application of AI in the patient's management. This ranges from explaining how AI will be applied, the data to be captured, and potential consequences on their diagnosis and treatment. Legal issues also extend to issues of legal responsibility where there is a question of whether an AI assisted diagnosis was right or end results were unfavourable. Healthcare managers must consult the legal advisors that provide them with guidelines and expectations of AI application in practice fields.

5.4. Limitations of the Study

Irrespective of the identified study's contributions and implications, it had some limitations. These limitations didn't hamper the significant findings; they just highlighted some deficiencies the researcher observed during the whole research journey. The limitations have different types as methodological, data limitation and context limitations. In this myriad, the study has its first limitation as the methodological limitation in which it struggled with data deficiency. The visual data used for analysis had different codes and classes but some images were not labeled and, therefore excluded from the data analysis.

Due to this minority classes discrepancy, the study had to face a significant loss and this may have caused the scope and dispersion of the interpretated results. In total, there were 18082 images with no issue, and the rest of the images counted only 1918 which had some useful labels reflective of the designed categories. This limited data availability caused different challenges in data analysis due to the decreased accuracy and precision of the models. In this manner, the integrated deep learning model algorithms such as SVM and CNN faced some struggling points and denoted slightly lower precision and the SVM algorithm encountered some laggings in its functioning in all designed categories. Thus, the data limitation was a significant issue that created many struggling points for the implemented algorithms to perform and added limitations to their functioning.

In addition to the data set-related limitation, this study also holds some methodological limitations associated with the models used for breast cancer detection. This study has used AI-based deep learning as four different algorithms, but, the era of massive advancement has taken another update as the integration of virtual reality and augmented reality type AI technologies. This integration has been gaining recognition due to its potential to bolster patient engagement, enhance surgical precision, and customize medical interventions [185]. But this study only used the primary stage or level of AI technology which adds to its limitations.

Next, the study has its contextual limitation as it has only emphasized breast cancer and its associated deep learning detection and treatment models. There is no doubt that breast cancer is an alleviating health concern and it has been gaining massive steam in recent years in both genders therefore, this study has effectively presented different models for its early and accurate detection. Still, there are other types of cancers as well that have more dreading and devastating outcomes on the human body which have to be investigated and their origins also underscore significant attention [186]. However, this study was confined to only breast cancer and it has no or least implications for the detection systems, models and techniques for detection.

In addition to this, this study has several other limitations as interpretability, generalizability and clinical validations. The study has followed a simplistic approach and has used the deep learning model without providing any transparency on the changes fostered in the model predictors which have a significant value in clinical research. In addition to the limited data sample, the medical history and other personal information were not properly monitored which left a question mark on the quality of the data. In the last, the authors used efficient resources for running the algorithms and kept the best standard in view but still limited, and the deficiency of essentially stored data was due to the limitations in the data resources as well. This study has used efficient software with adequate GPUs, but still, if the data set would be large, then the study needed more powerful software and resources to ensure quality data storage and algorithm processing.

Next, the data quality and quantity both had to be compromised which caused another concern of reservations about the ethicality of the used data set in the research. The basic reason behind this limitation was the lack of powerful and highly rigid data sources. The study used visual graphics as the raw data to perform the selected algorithms and reported the derived results in this study, but have mixed perceptions regarding data privacy and ethical concerns. Normally, in studies, the visual or graphical plotting and reporting of personal information is prohibited [187], but this study has gathered the secondary data from a website and it's not a confirmed assumption that the visual data shared there has been uploaded with the considerations of ethical concerns or not. Thus, all of these limitations and most importantly the deficit data sources limitations were the major observed limitations that have caused significant other limitations in the incorporated algorithms and their accuracies as well.

5.5. Future Research Directions

To fill the identified deficiencies in the study, future researchers and empirical scholars can use the subsequent discussed research directions to uncover more fascinating facts about the conjunction of AI technology and cancer diagnosis and treatments. In the future, the potential studies can extend the scope of this study by using a large data set comprised of at least more than 250 images or maybe more than this or visual presentations to deeply analyze the deep learning models with diverse types of algorithms and their accuracy. In addition, future studies have a chance to find more accurate and precise results against the integrated algorithms. A large sample set will be a great opportunity to identify the accuracy and precision of the deep learning model with different algorithms similar to or other than those that were used in this study and share complementing or enhanced findings to the literature.

Next, future studies can use the combined algorithms of basic deep learning models and AI visual technology such as visual reality VR and augmented reality AR to explore breast cancer from a deeper and holistic view. The studies can expand their methodological approach by integrating the deep learning models to identify suspicious and malicious cases and then can use the AI visual technologies to take a 3D view of the suspected cases for more accurate precision and diagnosis, and inexpensive diagnosis or data collection practices [188]. With this strategy, future studies can perform a deeper investigation to identify the coherence between different AI technologies. Moreover, the potential scholars can extend the scope of this study by integrating more deep learning models to identify the most accurate model in deep learning for breast cancer detection.

Furthermore, scholars can explore cancer diseases in other organs such as the liver, esophagus, pancreas, kidneys, etc. With this approach, empirical scholars will grab the opportunity to use a novel idea to use AI technology in other organs, but studies will require advanced and basic AI technologies for a better strategic approach. In addition to this, future studies can use different nature data sets and can gather fresh data using an experimental method for research. Moreover, the studies can use a comparative strategy in which the conventional diagnosis methods can be compared with the advanced deep learning or virtual reality style diagnosis methods to provide a deeper understanding of the functions of conventional and advanced technologies. With this comparison, the studies can facilitate scholars and practical workers to have a deeper understanding of the accuracy and precision of cancer-diagnosis technologies to explore and move ahead in the field of technology. This study has a limitation in the source of data which was secondary to this research and it caused different reservations in the incorporated strategies as well.

Therefore, it is highly suggested to future scholars to use primary data with the prior consent of the patients to eliminate the chance of any discrepancy in the data set and the reservation related to ethical concerns and privacy. Data privacy is one of the most fundamental concepts in research ethics, therefore, it is highly recommended to use first-hand data collected from the participants' consent with prelude information sharing about the use of data. With this practice, the ethicality of the research will be not compromised.

Moreover, this study was restricted to only one deep learning model, so next studies can use the opportunity of using two or more deep learning models with different algorithms. Moreover, the next studies can extend this same study by integrating the missing elements of interpretability, generalizability and clinical validations. With these components, one can enhance the transparency of the model with a comprehensive explanation of the changes in the model predictors. Moreover, with clinical validations, the studies will be able to add the component of practical justifications in the identified result and will be a piece of solid evidence in the literature.

In the last, it is a bit of strong advice for the next studies to keep a powerful and substantial resource compatible to handle the load of the software for algorithms' efficient working. Because, if one wants to use the same idea of this study, he or she must adhere to a large sample set that has to be stored in stable and high-memory storage devices so that the quality of data can't be compromised. In the end, future researchers can have a free space to explore the literature and practical field relevant to deep learning and breast or any other organ cancer with novel ideation to produce meaningful insights for the empirical and practical field.

Chapter 6: Conclusion

This thesis explored the application of deep learning techniques which are enhancing the detection of breast cancer by focusing on the development and evaluation of advanced models which addresses the limitations of the traditional imaging and computer-aided diagnostics (CAD) systems. The aim of the research is to improve the accuracy, reliability and efficiency of diagnosis methods used for breast cancer and also contributes to improve the outcomes of patients and to provide them with effective clinical results. The research also facilitates the clinical practices as it provides them with a model for effective diagnosis of breast cancer. The study was motivated by the extensive need of healthcare department as they require advance breast cancer detection methodologies. The conventional imaging techniques are beneficial in the field such as the mammography, ultrasound, and MRI but these techniques have limitations regarding the sensitivity, specificity, and accuracy. The aim of the research is to address these limitations by using the deep learning models and techniques which shows promising results in different domains which also involves medical imaging.

The methodology chapter lays the groundwork for the study as it provides with detailed processes which are used for data preparation and development. The INbreast database provides with comprehensive dataset for the training and validation of the models. This process involves the denoising and removing of pectoral muscles from the mammogram images which enhances the image quality and ensures that the models receive high-quality inputs. These preprocessing techniques are crucial for the improving the effectiveness of the deep learning models. For addressing the limited size of the dataset, data augmentation techniques were applied. These techniques increase the diversity and volume of the training data which is necessary for the improvement of model generalization and performance. The research is focused on advanced deep learning architectures such as YOLOv8 and Mask R-CNN. The YOLOv8 is used for the object detection capabilities while the Mask R-CNN is applied for the segmentation and tumor size prediction features. Both the models are modified and improvised in order to enhance their performance in detecting and diagnosing breast cancer. The model evaluation is a critical aspect of the study as various metrics were used for assessing their performance. The model tested involves Random Forest which emerges as the most effective model for the detection of breast cancer. The model demonstrated high accuracy across different BI-RADS categories such as BI-RAD-3, BI-RAD-4, BI-RAD-5 and non-labeled images. The model was reliable as it distinguishes between different levels of suspicion and malignancy. The robustness and accuracy make the model a valuable tool for clinical diagnosis and decision-making. The Support Vector Machine (SVM) shows the potential in the creation of clear decision boundaries but also requires further tuning for achieving balanced performance across all BI-RADS categories. The effectiveness improves the additional parameter optimization and refinement. The Convolutional Neural Networks (CNNs) have been successful in image analysis tasks, but the model faces challenges in the study because of data limitations and model parameter adjustments. The results of the findings indicate that there is need for more extensive training of data in order to detect the breast cancer effectively.

The findings of the research have significant implications for clinical practices as it suggests that Random Forest model's high accuracy in identifying various BI-RADS categories supports the reliability in breast cancer detection. This improvement in the diagnostic accuracy is essential for the early detection and treatment which leads to the improvement of patient outcomes and increases their survival rates. By minimizing the false positives and false negatives the models enhance the diagnostic precisions. This reduction in errors decreases the risk of missed diagnosis and the unnecessary biopsies which ultimately reduces the patient stress and improves the overall efficiency of diagnostics. By integrating these advanced models into the clinical workflow, streamlines the diagnostic process. The ability to automate the initial screening tasks allows radiologists to focus on more complex cases by alleviating the workload and reducing the fatigue. The efficiency leads to faster diagnosis and treatment which provides benefit to both the patient and healthcare providers. The research contributes to the field of breast cancer as it facilitates the detection of cancer in several ways. The application of YOLOv8 and Mask R-CNN in the breast cancer detection demonstrates the effectiveness of deep learning as it addresses the complex challenges in the medical imaging. These models offer with significant improvements over the traditional methods and provides foundation to the future advancements in the field. The development and use of Random Forest model achieves high accuracy and reliability and offers with practical solutions for the enhancement of breast cancer detection. The model can also be integrated into the clinical practices as it supports more accurate and efficient diagnosis. The comparative analysis of the different models provides with valuable insights into the strengths and limitations. The study guides the future research and development efforts as it helps in refinement and optimization of deep learning techniques for the detection of breast cancer.

The thesis successfully explains the potentials of deep learning models for advancing the field of breast cancer detection. By using the advanced techniques and models, the research addresses the key limitations of the traditional diagnosis methods such as CAD systems. The Random Forest model has proven to be the most effective model as it provides with accurate, reliable, and efficient diagnosis. The study has some limitations and challenges which are associated with the quality of data, performance of the model and its implementation but the contributions are significant. The findings highlight the importance of deep learning as it improves the diagnosis of breast cancer and provides with solid foundations for integration of technology into the clinical diagnostic practices. The achieved advancements of the research have potential to enhance the patient outcomes and to streamline the diagnostic workflow in order to contribute to the ongoing efforts of breast cancer diagnosis. The thesis significantly highlights the transformative impact of deep learning in the detection of breast cancer and emphasized the importance of innovation and development in the medical imaging area.

References

1. Mustafa, M., et al., *Molecular pathways and therapeutic targets linked to triple-negative breast cancer (TNBC).* Molecular and Cellular Biochemistry, 2024. **479**(4): p. 895-913.
2. Organization, W.H., *Global breast cancer initiative implementation framework: assessing, strengthening and scaling-up of services for the early detection and management of breast cancer*. 2023: World Health Organization.
3. Taylor, C., et al., *Breast cancer mortality in 500 000 women with early invasive breast cancer diagnosed in England, 1993-2015: population based observational cohort study.* Bmj, 2023. **381**.
4. Nida, B.A., T.B. Rooney, and M.M. Miller, *Utility of MRI-directed contrast-enhanced mammography for biopsy planning for suspicious MRI-detected breast lesions.* American Journal of Roentgenology, 2023. **220**(2): p. 202-211.
5. Teklić, J. and T. Matijaš, *Comparison between breast magnetic resonance imaging and contrast-enhanced mammography.* Hrvatski časopis zdravstvenih znanosti, 2023. **2**(3): p. 108-119.
6. Najjar, R., *Redefining radiology: a review of artificial intelligence integration in medical imaging.* Diagnostics, 2023. **13**(17): p. 2760.
7. Li, M., et al., *Medical image analysis using deep learning algorithms.* Frontiers in Public Health, 2023. **11**: p. 1273253.
8. Sharmin, S., et al., *A hybrid dependable deep feature extraction and ensemble-based machine learning approach for breast cancer detection.* IEEE Access, 2023.
9. Jafari, Z. and E. Karami, *Breast cancer detection in mammography images: a CNN-based approach with feature selection.* Information, 2023. **14**(7): p. 410.
10. Albawi, S., M.H. Arif, and J. Waleed, *Skin cancer classification dermatologist-level based on deep learning model.* Acta Scientiarum. Technology, 2023. **45**: p. e61531-e61531.
11. Abhisheka, B., S.K. Biswas, and B. Purkayastha, *A comprehensive review on breast cancer detection, classification and segmentation using deep learning.* Archives of Computational Methods in Engineering, 2023. **30**(8): p. 5023-5052.
12. Zakareya, S., H. Izadkhah, and J. Karimpour, *A new deep-learning-based model for breast cancer diagnosis from medical images.* Diagnostics, 2023. **13**(11): p. 1944.
13. Neha Thakur, A.K., Pardeep Kumar, *A systematic review of machine and deep learning techniques for the identification and classification of breast cancer through medical image modalities.* 2023.
14. Dou, B., et al., *Machine learning methods for small data challenges in molecular science.* Chemical Reviews, 2023. **123**(13): p. 8736-8780.
15. Dingle, L., *A comparison and analysis of explainable clinical decision making using white box and black box models*. 2024.
16. Borys, K., et al., *Explainable AI in medical imaging: An overview for clinical practitioners–Beyond saliency-based XAI approaches.* European journal of radiology, 2023. **162**: p. 110786.
17. Aggarwal, K., et al., *Developing and deploying deep learning models in brain magnetic resonance imaging: A review.* NMR in Biomedicine, 2023. **36**(12): p. e5014.

18. Alam, M.N., M. Kaur, and M.S. Kabir, *Explainable AI in Healthcare: Enhancing transparency and trust upon legal and ethical consideration.* Int Res J Eng Technol, 2023. **10**(6): p. 1-9.

19. Yusoff, M., et al., *Accuracy analysis of deep learning methods in breast cancer classification: A structured review.* Diagnostics, 2023. **13**(4): p. 683.

20. Salehi, A.W., et al., *A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope.* Sustainability, 2023. **15**(7): p. 5930.

21. Wang, Z., et al., *Deep learning-based multi-modal data integration enhancing breast cancer disease-free survival prediction.* Precision Clinical Medicine, 2024. **7**(2).

22. Ahn, J.S., et al., *Artificial intelligence in breast cancer diagnosis and personalized medicine.* Journal of Breast Cancer, 2023. **26**(5): p. 405.

23. Taylor, C.R., et al., *Artificial intelligence applications in breast imaging: current status and future directions.* Diagnostics, 2023. **13**(12): p. 2041.

24. Najjar, R., *Digital Frontiers in Healthcare: Integrating mHealth, AI, and Radiology for Future Medical Diagnostics.* 2024.

25. Nasser, M. and U.K. Yusof, *Deep learning based methods for breast cancer diagnosis: a systematic review and future direction.* Diagnostics, 2023. **13**(1): p. 161.

26. Gong, Y., et al., *A survey on dataset quality in machine learning.* Information and Software Technology, 2023. **162**: p. 107268.

27. Logan, J., P.J. Kennedy, and D. Catchpoole, *A review of the machine learning datasets in mammography, their adherence to the FAIR principles and the outlook for the future.* Scientific Data, 2023. **10**(1): p. 595.

28. Ben Ammar, M., et al., *Harnessing Deep Learning for Early Breast Cancer Diagnosis: A Review of Datasets, Methods, Challenges, and Future Directions.* International Journal of Computing and Digital Systems, 2024. **15**(1): p. 1-13.

29. Kurt, S.A., et al., *Comparing the efficiency of imaging modalities in detection of recurrent breast cancer.* European Journal of Breast Health, 2023. **19**(1): p. 85.

30. Shaik, T., et al., *Multimodality Fusion for Smart Healthcare: a Journey from Data, Information, Knowledge to Wisdom.* arXiv preprint arXiv:2306.11963, 2023.

31. Adam, R., et al., *Deep learning applications to breast cancer detection by magnetic resonance imaging: a literature review.* Breast Cancer Research, 2023. **25**(1): p. 87.

32. Loizidou, K., R. Elia, and C. Pitris, *Computer-aided breast cancer detection and classification in mammography: A comprehensive review.* Computers in Biology and Medicine, 2023. **153**: p. 106554.

33. Zhang, Q., et al., *Multimodal fusion on low-quality data: A comprehensive survey.* arXiv preprint arXiv:2404.18947, 2024.

34. Paudyal, R., et al., *Artificial intelligence in CT and MR imaging for oncological applications.* Cancers, 2023. **15**(9): p. 2573.

35. Ahmed, M., et al., *Enhancing Breast Cancer Diagnosis in Mammography: Evaluation and Integration of Convolutional Neural Networks and Explainable AI.* arXiv preprint arXiv:2404.03892, 2024.

36. Vázquez-Lema, D., et al., *Segmentation, Classification and Interpretation of Breast Cancer Medical Images using Human-in-the-Loop Machine Learning.* arXiv preprint arXiv:2403.20112, 2024.

37. Xu, Q., et al., *Interpretability of clinical decision support systems based on artificial intelligence from technological and medical perspective: A systematic review.* Journal of healthcare engineering, 2023. **2023**(1): p. 9919269.

38. Knevel, R. and K.P. Liao, *From real-world electronic health record data to real-world results using artificial intelligence.* Annals of the Rheumatic Diseases, 2023. **82**(3): p. 306-311.

39. Bartoldson, B.R., B. Kailkhura, and D. Blalock, *Compute-efficient deep learning: Algorithmic trends and opportunities.* Journal of Machine Learning Research, 2023. **24**(122): p. 1-77.

40. Zhai, K., et al., *Optimizing clinical workflow using precision medicine and advanced data analytics.* Processes, 2023. **11**(3): p. 939.

41. Balch, J.A., et al., *Machine learning–enabled clinical information systems using fast healthcare interoperability resources data standards: scoping review.* JMIR Medical Informatics, 2023. **11**: p. e48297.

42. Xue, P., et al., *Unassisted Clinicians Versus Deep Learning–Assisted Clinicians in Image-Based Cancer Diagnostics: Systematic Review With Meta-analysis.* Journal of Medical Internet Research, 2023. **25**: p. e43832.

43. Franklin, G., et al., *The Sociodemographic Biases in Machine Learning Algorithms: A Biomedical Informatics Perspective.* Life, 2024. **14**(6): p. 652.

44. Hort, M., et al., *Bias mitigation for machine learning classifiers: A comprehensive survey.* ACM Journal on Responsible Computing, 2024. **1**(2): p. 1-52.

45. Zhang, Y.-P., et al., *Artificial intelligence-driven radiomics study in cancer: the role of feature engineering and modeling.* Military Medical Research, 2023. **10**(1): p. 22.

46. Jan, Y.-T., et al., *Machine learning combined with radiomics and deep learning features extracted from CT images: a novel AI model to distinguish benign from malignant ovarian tumors.* Insights into Imaging, 2023. **14**(1): p. 68.

47. Wei, L., et al., *Artificial intelligence (AI) and machine learning (ML) in precision oncology: a review on enhancing discoverability through multiomics integration.* The British Journal of Radiology, 2023. **96**(1150): p. 20230211.

48. Javed, H., et al., *Ethical Frameworks for Machine Learning in Sensitive Healthcare Applications.* IEEE Access, 2023.

49. Nadella, G.S., et al., *A Systematic Literature Review of Advancements, Challenges and Future Directions of AI And ML in Healthcare.* International Journal of Machine Learning for Sustainable Development, 2023. **5**(3): p. 115-130.

50. Kim, J.P., et al., *Physicians' and machine learning researchers' perspectives on ethical issues in the early development of clinical machine learning tools: qualitative interview study.* JMIR AI, 2023. **2**: p. e47449.

51. Luo, L., et al., *Deep learning in breast cancer imaging: A decade of progress and future directions.* IEEE Reviews in Biomedical Engineering, 2024.

52. Illimoottil, M. and D. Ginat, *Recent advances in deep learning and medical imaging for head and neck cancer treatment: MRI, CT, and PET scans.* Cancers, 2023. **15**(13): p. 3267.

53. Karabacak, M. and K. Margetis, *Embracing large language models for medical applications: opportunities and challenges.* Cureus, 2023. **15**(5).

54. Hu, Q., H.M. Whitney, and M.L. Giger, *A deep learning methodology for improved breast cancer diagnosis using multiparametric MRI.* Scientific reports, 2020. **10**(1): p. 10536.

55. Mridha, M.F., et al., *A comprehensive survey on deep-learning-based breast cancer diagnosis.* Cancers, 2021. **13**(23): p. 6116.

56. Lou, B., et al., *An image-based deep learning framework for individualising radiotherapy dose: a retrospective analysis of outcome prediction.* The Lancet Digital Health, 2019. **1**(3): p. e136-e147.

57. Parikh, J.R., J. Sun, and M.B. Mainiero, *What causes the most stress in breast radiology practice? A survey of members of the Society of Breast Imaging.* Journal of breast imaging, 2021. **3**(3): p. 332-342.

58. Lotter, W., et al., *Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach.* Nature medicine, 2021. **27**(2): p. 244-249.

59. Onchong'a, A.G., *Demand, Perceived Barriers and Available Infrastructure for Radiology Subspecialisation Training in Kenya*. 2022, University of Nairobi.

60. Krones, F.H. and B. Walker, *From theoretical models to practical deployment: A perspective and case study of opportunities and challenges in AI-driven healthcare research for low-income settings.* medRxiv, 2023: p. 2023.12. 26.23300539.

61. Alzubaidi, L., et al., *Review of deep learning: concepts, CNN architectures, challenges, applications, future directions.* Journal of big Data, 2021. **8**: p. 1-74.

62. Alowais, S.A., et al., *Revolutionizing healthcare: the role of artificial intelligence in clinical practice.* BMC medical education, 2023. **23**(1): p. 689.

63. Huang, G., et al., *A systematic review and meta-analysis of diagnostic performance and physicians’ perceptions of artificial intelligence (AI)-assisted CT diagnostic technology for the classification of pulmonary nodules.* Journal of Thoracic Disease, 2021. **13**(8): p. 4797.

64. Castiglioni, I., et al., *AI applications to medical images: From machine learning to deep learning.* Physica medica, 2021. **83**: p. 9-24.

65. Prathyusha, T., C. Vallepalli, and K. Nagaraj, *Knowledge on Risk Factors and Early Detection Methods of Breast Cancer among Rural Women of Chittoor District, Andhra Pradesh, India.* 2022.

66. Jaglan, P., R. Dass, and M. Duhan, *Breast cancer detection techniques: issues and challenges.* Journal of The Institution of Engineers (India): Series B, 2019. **100**(4): p. 379-386.

67. Liu, X., et al., *A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis.* The lancet digital health, 2019. **1**(6): p. e271-e297.

68. Lawrence, L.C., *Machine Learning Integration in Healthcare: Human Error Reduction in Breast Cancer Screening*. 2024, Colorado Technical University.

69. Hilabi, B.S., S.A. Alghamdi, and M. Almanaa, *Impact of magnetic resonance imaging on healthcare in low-and middle-income countries.* Cureus, 2023. **15**(4).

70. Kakileti, S.T., *Machine learning for breast cancer diagnosis in developing Countries.* 2020.

71. Bazoukis, G., et al., *The inclusion of augmented intelligence in medicine: a framework for successful implementation.* Cell Reports Medicine, 2022. **3**(1).

72. Maleki Varnosfaderani, S. and M. Forouzanfar, *The role of AI in hospitals and clinics: transforming healthcare in the 21st century.* Bioengineering, 2024. **11**(4): p. 337.

73. Echle, A., et al., *Deep learning in cancer pathology: a new generation of clinical biomarkers.* British journal of cancer, 2021. **124**(4): p. 686-696.

74. Sung, H., et al., *Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA: a cancer journal for clinicians, 2021. **71**(3): p. 209-249.

75. Crosby, D., et al., *Early detection of cancer.* Science, 2022. **375**(6586): p. eaay9040.

76. Michael, E., et al., *Breast cancer segmentation methods: current status and future potentials.* BioMed research international, 2021. **2021**(1): p. 9962109.

77. Zubair, M., S. Wang, and N. Ali, *Advanced approaches to breast cancer classification and diagnosis.* Frontiers in Pharmacology, 2021. **11**: p. 632079.

78. Bilani, N., et al., *Breast cancer in the United States: A Cross-Sectional overview.* Journal of cancer epidemiology, 2020. **2020**(1): p. 6387378.

79. Bhattacharyya, G.S., et al., *Overview of breast cancer and implications of overtreatment of early-stage breast cancer: an Indian perspective.* JCO global oncology, 2020. **6**: p. 789-798.

80. Wilkinson, L. and T. Gathani, *Understanding breast cancer as a global health concern.* The British journal of radiology, 2022. **95**(1130): p. 20211033.

81. Wang, J. and S.-G. Wu, *Breast cancer: an overview of current therapeutic strategies, challenge, and perspectives.* Breast Cancer: Targets and Therapy, 2023: p. 721-730.

82. Whitaker, K., *Earlier diagnosis: the importance of cancer symptoms.* The Lancet Oncology, 2020. **21**(1): p. 6-8.

83. Aziz, M.H. and A.A. Abdulla, *Computer-aided Diagnosis for the Early Breast Cancer Detection.* UHD Journal of Science and Technology, 2023. **7**(1): p. 7-14.

84. Baccouche, A., *Computer aided diagnosis system for breast cancer using deep learning.* 2022.

85. Guo, Z., et al., *A review of the current state of the computer-aided diagnosis (CAD) systems for breast cancer diagnosis.* Open Life Sciences, 2022. **17**(1): p. 1600-1611.

86. van Zelst, J.C., et al., *Validation of radiologists' findings by computer-aided detection (CAD) software in breast cancer detection with automated 3D breast ultrasound: a concept study in implementation of artificial intelligence software.* Acta Radiologica, 2020. **61**(3): p. 312-320.

87. Ramadan, S.Z., *Methods Used in Computer-Aided Diagnosis for Breast Cancer Detection Using Mammograms: A Review.* Journal of healthcare engineering, 2020. **2020**(1): p. 9162464.

88. Arun Kumar, S. and S. Sasikala, *Review on deep learning-based CAD systems for breast cancer diagnosis.* Technology in Cancer Research & Treatment, 2023. **22**: p. 15330338231177977.

89. Eltrass, A.S. and M.S. Salama, *Fully automated scheme for computer-aided detection and breast cancer diagnosis using digitised mammograms.* IET Image Processing, 2020. **14**(3): p. 495-505.

90. Bratinčević, L. and T. Matijaš, *Application of CAD in the diagnosis of breast cancer.* Radiology News Journal/Radiološki Vjesnik, 2022. **46**(1).

91. Dar, R.A., M. Rasool, and A. Assad, *Breast cancer detection using deep learning: Datasets, methods, and challenges ahead.* Computers in biology and medicine, 2022. **149**: p. 106073.

92. Houssein, E.H., et al., *Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review.* Expert Systems with Applications, 2021. **167**: p. 114161.

93. Saber, A., et al., *A novel deep-learning model for automatic detection and classification of breast cancer using the transfer-learning technique.* IEEe Access, 2021. **9**: p. 71194-71209.

94. Salvi, S. and A. Kadam. *Breast Cancer Detection Using Deep learning and IoT Technologies*. in *Journal of Physics: Conference Series*. 2021. IOP Publishing.

95. Allugunti, V.R., *Breast cancer detection based on thermographic images using machine learning and deep learning algorithms.* International Journal of Engineering in Computer Science, 2022. **4**(1): p. 49-56.

96. Alanazi, S.A., et al., *Boosting breast cancer detection using convolutional neural network.* 2021. **2021**(1): p. 5528622.

97. Masud, M., et al., *Convolutional neural network-based models for diagnosis of breast cancer.* 2022. **34**(14): p. 11383-11394.

98. Heenaye-Mamode Khan, M., et al., *Multi-class classification of breast cancer abnormalities using Deep Convolutional Neural Network (CNN).* 2021. **16**(8): p. e0256500.

99. Yurttakal, A.H., et al., *Detection of breast cancer via deep convolution neural networks using MRI images.* 2020. **79**(21): p. 15555-15573.

100. Sood, A.J.N.J.f.A.I. and I.o. Things, *Breast cancer detection using neural networks.* 2023. **1**(1): p. 12-18.

101. Ereken, Ö.F. and C. Tarhan. *Breast cancer detection using convolutional neural networks*. in *2022 International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. 2022. IEEE.

102. Desai, M. and M.J.C.e. Shah, *An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN).* 2021. **4**: p. 1-11.

103. Patil, R.S. and N.J.E.i. Biradar, *Automated mammogram breast cancer detection using the optimized combination of convolutional and recurrent neural network.* 2021. **14**(4): p. 1459-1474.

104. Yao, H., et al., *Parallel structure deep neural network using CNN and RNN with an attention mechanism for breast cancer histology image classification.* 2019. **11**(12): p. 1901.

105. Abdul Hayum, A., et al., *A modified recurrent neural network (MRNN) model for and breast cancer classification system.* 2023. **64**(4): p. 1193-1203.

106. Soni, K.M., A. Gupta, and T. Jain. *Supervised machine learning approaches for breast cancer classification and a high performance recurrent neural network*. in *2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)*. 2021. IEEE.

107. Saleh, H., et al., *Predicting breast cancer based on optimized deep learning approach.* 2022. **2022**(1): p. 1820777.

108. Dutta, S., et al., *Breast cancer prediction using stacked GRU-LSTM-BRNN.* 2020. **25**(2): p. 163-171.

109. Ebrahim, M., A.A.H. Sedky, and S.J.D. Mesbah, *Accuracy assessment of machine learning algorithms used to predict breast cancer.* 2023. **8**(2): p. 35.

110. Das, A., M.N.J.A.C. Mohanty, and Informatics, *Design of ensemble recurrent model with stacked fuzzy ARTMAP for breast cancer detection.* 2022.

111. Chaudhury, S. and K.J.H.A. Sau, *A blockchain-enabled internet of medical things system for breast cancer detection in healthcare.* 2023. **4**: p. 100221.

112. Talasila, V., et al., *The Prediction of Diseases Using Rough Set Theory with Recurrent Neural Network in Big Data Analytics.* 2020. **13**(5).

113. Ayana, G., K. Dese, and S.-w.J.C. Choe, *Transfer learning in breast cancer diagnoses via ultrasound imaging.* 2021. **13**(4): p. 738.

114. Alzubaidi, L., et al., *Optimizing the performance of breast cancer classification by employing the same domain transfer learning from hybrid deep convolutional neural network model.* 2020. **9**(3): p. 445.

115. Alruwaili, M. and W.J.S. Gouda, *Automated breast cancer detection models based on transfer learning.* 2022. **22**(3): p. 876.

116. Saber, A., et al., *A novel deep-learning model for automatic detection and classification of breast cancer using the transfer-learning technique.* 2021. **9**: p. 71194-71209.

117. Nasir, M.U., et al., *Breast Cancer Prediction Empowered with Fine-Tuning.* 2022. **2022**(1): p. 5918686.

118. Vrbančič, G. and V.J.I.A. Podgorelec, *Transfer learning with adaptive fine-tuning.* 2020. **8**: p. 196197-196211.

119. Chen, Y., et al. *Fine-tuning ResNet for breast cancer classification from mammography*. in *Proceedings of the 2nd International Conference on Healthcare Science and Engineering 2nd*. 2019. Springer.

120. Madani, M., M.M. Behzadi, and S. Nabavi, *The role of deep learning in advancing breast cancer detection using different imaging modalities: a systematic review.* Cancers, 2022. **14**(21): p. 5334.

121. Aristokli, N., et al., *Comparison of the diagnostic performance of Magnetic Resonance Imaging (MRI), ultrasound and mammography for detection of breast cancer based on tumor type, breast density and patient's history: A review.* Radiography, 2022. **28**(3): p. 848-856.

122. Leong, Y.S., et al., *Microcalcification discrimination in mammography using deep convolutional neural network: towards rapid and early breast cancer diagnosis.* Frontiers in public health, 2022. **10**: p. 875305.

123. Bhowmik, A. and S. Eskreis-Winkler, *Deep learning in breast imaging.* BJR| Open, 2022. **4**(1): p. 20210060.

124. Eloranta, S. and M. Boman, *Predictive models for clinical decision making: Deep dives in practical machine learning.* Journal of Internal Medicine, 2022. **292**(2): p. 278-295.

125. Abdou, M.A., *Literature review: Efficient deep neural networks techniques for medical image analysis.* Neural Computing and Applications, 2022. **34**(8): p. 5791-5812.

126. Gomathi, P., C. Muniraj, and P. Periasamy, *Micro Calcification Detection in Mammogram Images Using Contiguous Convolutional Neural Network Algorithm.* Computer Systems Science & Engineering, 2023. **45**(2).

127. Hassan, N.M., S. Hamad, and K. Mahar, *Mammogram breast cancer CAD systems for mass detection and classification: a review.* Multimedia Tools and Applications, 2022. **81**(14): p. 20043-20075.

128. Zhang, J., et al. *Recent advancements in artificial intelligence for breast cancer: Image augmentation, segmentation, diagnosis, and prognosis approaches*. in *Seminars in Cancer Biology*. 2023. Elsevier.

129. Fiorentino, M.C., et al., *A review on deep-learning algorithms for fetal ultrasound-image analysis.* Medical image analysis, 2023. **83**: p. 102629.

130. de Santana Correia, A. and E.L. Colombini, *Attention, please! A survey of neural attention models in deep learning.* Artificial Intelligence Review, 2022. **55**(8): p. 6037-6124.

131. Tsuneki, M., *Deep learning models in medical image analysis.* Journal of Oral Biosciences, 2022. **64**(3): p. 312-320.

132. Balasubramaniam, S., et al., *A modified LeNet CNN for breast cancer diagnosis in ultrasound images.* Diagnostics, 2023. **13**(17): p. 2746.

133. Song, K., J. Feng, and D. Chen, *A survey on deep learning in medical ultrasound imaging.* Frontiers in Physics, 2024. **12**: p. 1398393.

134. Alsaadi, B.S.O., et al., *Hipaa Compliance In Medical Record Management: Ensuring Patient Privacy And Data Security. An Evolution.* Journal of Namibian Studies: History Politics Culture, 2022. **32**: p. 2585-2591.

135. Rahmat, K., et al., *MRI breast: current imaging trends, clinical applications, and future research directions.* Current Medical Imaging, 2022. **18**(13): p. 1347-1361.

136. Zhao, X., et al., *Clinical applications of deep learning in breast MRI.* Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 2023. **1878**(2): p. 188864.

137. Naga Srinivasu, P., et al., *Variational Autoencoders-BasedSelf-Learning Model for Tumor Identification and Impact Analysis from 2-D MRI Images.* Journal of Healthcare Engineering, 2023. **2023**(1): p. 1566123.

138. Ehrhardt, J. and M. Wilms, *Autoencoders and variational autoencoders in medical image analysis*, in *Biomedical Image Synthesis and Simulation*. 2022, Elsevier. p. 129-162.

139. Abunadi, I., *Deep and hybrid learning of MRI diagnosis for early detection of the progression stages in Alzheimer's disease.* Connection Science, 2022. **34**(1): p. 2395-2430.

140. Atasever, S., et al., *A comprehensive survey of deep learning research on medical image analysis with focus on transfer learning.* Clinical imaging, 2023. **94**: p. 18-41.

141. Lee, Y.-H., et al., *Advantages of deep learning with convolutional neural network in detecting disc displacement of the temporomandibular joint in magnetic resonance imaging.* Scientific Reports, 2022. **12**(1): p. 11352.

142. Mushtaq, Z., et al., *Effective K-nearest neighbor classifications for Wisconsin breast cancer data sets.* 2020. **43**(1): p. 80-92.

143. Jabbar, M.A.J.E. and A.S. Research, *Breast cancer data classification using ensemble machine learning.* 2021. **48**(1).

144. Ul Haq, A., et al., *A novel integrated diagnosis method for breast cancer detection.* 2020. **38**(2): p. 2383-2398.

145. Alanazi, S.A., et al., *Boosting breast cancer detection using convolutional neural network.* Journal of Healthcare Engineering, 2021. **2021**(1): p. 5528622.

146. Charan, S., M.J. Khan, and K. Khurshid. *Breast cancer detection in mammograms using convolutional neural network*. in *2018 international conference on computing, mathematics and engineering technologies (iCoMET)*. 2018. IEEE.

147. He, J., et al., *ReLU deep neural networks and linear finite elements.* arXiv preprint arXiv:1807.03973, 2018.

148. Kebede, S.R., et al., *Dual view deep learning for enhanced breast cancer screening using mammography.* Scientific Reports, 2024. **14**(1): p. 3839.

149. Liu, Z., et al. *Dropout reduces underfitting*. in *International Conference on Machine Learning*. 2023. PMLR.

150. Masud, M., A.E. Eldin Rashed, and M.S. Hossain, *Convolutional neural network-based models for diagnosis of breast cancer.* Neural Computing and Applications, 2022. **34**(14): p. 11383-11394.

151. Nagi, J., et al. *Max-pooling convolutional neural networks for vision-based hand gesture recognition*. in *2011 IEEE international conference on signal and image processing applications (ICSIPA)*. 2011. IEEE.

152. Nguyen, H.T., et al., *VinDr-Mammo: A large-scale benchmark dataset for computer-aided diagnosis in full-field digital mammography.* Scientific Data, 2023. **10**(1): p. 277.

153. Rajakumari, R. and L. Kalaivani, *Breast Cancer Detection and Classification Using Deep CNN Techniques.* Intelligent Automation & Soft Computing, 2022. **32**(2).

154. Wu, X., et al., *AAR-RT–a system for auto-contouring organs at risk on CT images for radiation therapy planning: principles, design, and large-scale evaluation on head-and-neck and thoracic cancer cases.* Medical image analysis, 2019. **54**: p. 45-62.

155. Wang, Q., et al., *Remote sensing image gap filling based on spatial-spectral random forests.* Science of Remote Sensing, 2022. **5**: p. 100048.

156. Tajbakhsh, N., et al., *Convolutional neural networks for medical image analysis: Full training or fine tuning?* IEEE transactions on medical imaging, 2016. **35**(5): p. 1299-1312.

157. Yasir, N., S. Anwar, and M.T. Khan, *Machine Vision based Intelligent Breast Cancer Detection.* Pakistan Journal of Engineering and Technology, 2022. **5**(1): p. 1-10.

158. Litjens, G., et al., *A survey on deep learning in medical image analysis.* Medical image analysis, 2017. **42**: p. 60-88.

159. Zhou, X., et al., *A comprehensive review for breast histopathology image analysis using classical and deep neural networks.* IEEE Access, 2020. **8**: p. 90931-90956.

160. Wang, K., et al., *A dual-mode deep transfer learning (D2TL) system for breast cancer detection using contrast enhanced digital mammograms.* IISE Transactions on Healthcare Systems Engineering, 2019. **9**(4): p. 357-370.

161. Kaul, D., H. Raju, and B. Tripathy, *Deep learning in healthcare.* Deep Learning in Data Analytics: Recent Techniques, Practices and Applications, 2022: p. 97-115.

162. Hanefi Calp, M., *Use of deep learning approaches in cancer diagnosis.* Deep learning for cancer diagnosis, 2021: p. 249-267.

163. Tran, K.A., et al., *Deep learning in cancer diagnosis, prognosis and treatment selection.* Genome Medicine, 2021. **13**: p. 1-17.

164. Chugh, G., S. Kumar, and N. Singh, *Survey on machine learning and deep learning applications in breast cancer diagnosis.* Cognitive Computation, 2021. **13**(6): p. 1451-1470.

165. Lin, R.-H., et al., *Application of deep learning to construct breast cancer diagnosis model.* Applied Sciences, 2022. **12**(4): p. 1957.

166. Akselrod-Ballin, A., et al., *Predicting breast cancer by applying deep learning to linked health records and mammograms.* Radiology, 2019. **292**(2): p. 331-342.

167. Houssami, N., et al., *Early detection of second breast cancers improves prognosis in breast cancer survivors.* Annals of oncology, 2009. **20**(9): p. 1505-1510.

168. Zheng, J., et al., *Deep learning assisted efficient AdaBoost algorithm for breast cancer detection and early diagnosis.* IEEE Access, 2020. **8**: p. 96946-96954.

169. Gopal, V.N., et al., *Feature selection and classification in breast cancer prediction using IoT and machine learning.* Measurement, 2021. **178**: p. 109442.

170. Mahesh, T., et al., *Early predictive model for breast cancer classification using blended ensemble learning.* International Journal of System Assurance Engineering and Management, 2024. **15**(1): p. 188-197.

171. Ebrahim, M., A.A.H. Sedky, and S. Mesbah, *Accuracy assessment of machine learning algorithms used to predict breast cancer.* Data, 2023. **8**(2): p. 35.

172. Zhao, Q. and Z. Shang. *Deep learning and its development*. in *Journal of Physics: Conference Series*. 2021. IOP Publishing.

173. Binder, A., et al., *Morphological and molecular breast cancer profiling through explainable machine learning.* Nature Machine Intelligence, 2021. **3**(4): p. 355-366.

174. Pereira, R.d.O., et al., *Evaluation of the accuracy of mammography, ultrasound and magnetic resonance imaging in suspect breast lesions.* Clinics, 2020. **75**: p. e1805.

175. Nicosia, L., et al. *History of mammography: analysis of breast imaging diagnostic achievements over the last century*. in *Healthcare*. 2023. MDPI.

176. Wu, X., et al., *Long short-term memory model–a deep learning approach for medical data with irregularity in cancer predication with tumor markers.* Computers in Biology and Medicine, 2022. **144**: p. 105362.

177. O'Keefe, K., *An exploration of radiographer decision-making regarding rejected or sub-optimal plain x-ray images*. 2023, Queensland University of Technology.

178. Hanna, T.P., et al., *Mortality due to cancer treatment delay: systematic review and meta-analysis.* bmj, 2020. **371**.

179. Ozcan, B.B., et al., *Breast cancer disparity and outcomes in underserved women.* RadioGraphics, 2023. **44**(1): p. e230090.

180. Bray, G.A. and D.H. Ryan, *Evidence-based weight loss interventions: individualized treatment options to maximize patient outcomes.* Diabetes, Obesity and Metabolism, 2021. **23**: p. 50-62.

181. Ajantha Devi, V. and A. Nayyar, *Fusion of deep learning and image processing techniques for breast cancer diagnosis.* Deep learning for cancer diagnosis, 2021: p. 1-25.

182. Lastrucci, A. and D. Giansanti, *Radiological Crossroads: Navigating the Intersection of Virtual Reality and Digital Radiology through a Comprehensive Narrative Review of Reviews.* Robotics, 2024. **13**(5): p. 69.

183. Oala, L., et al., *Machine learning for health: algorithm auditing & quality control.* Journal of medical systems, 2021. **45**: p. 1-8.

184. Yigzaw, K.Y., et al., *Health data security and privacy: Challenges and solutions for the future.* Roadmap to Successful Digital Health Ecosystems, 2022: p. 335-362.

185. Lastrucci, A., et al., *Exploring Augmented Reality Integration in Diagnostic Imaging: Myth or Reality?* Diagnostics, 2024. **14**(13): p. 1333.

186. Tzenios, N., *Obesity as a risk factor for cancer.* EPRA International Journal of Research and Development (IJRD), 2023. **8**(2): p. 101-104.

187. Avraam, D., et al., *Privacy preserving data visualizations.* EPJ data science, 2021. **10**(1): p. 2.

188. Lakshminarayanan, V., et al., *Health care equity through intelligent edge computing and augmented reality/virtual reality: a systematic review.* Journal of Multidisciplinary Healthcare, 2023: p. 2839-2859.