A Comprehensive Analysis of Machine Learning Algorithms for Breast Cancer Classification Across Diverse Image Modalities

Sania Azhmee Bhuiyan
Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
saniaazhmee98@gmail.com

Sabista Ifraj
Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
sbifraj123@gmail.com

Arpita Roy
Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
arpita.roy191203@gmail.com

Sohanoor Rahman
Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
sohanoorrahmansourav@gmail.com

Md Humaion Kabir Mehedi
Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
sohanoorrahmansourav@gmail.com

Annajiat Alim Rasel
Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
sohanoorrahmansourav@gmail.com

***Abstract*—This document is a model and instructions for LaTeX. This and the IEEEtran.cls file define the components of your paper [title, text, heads, etc.]. *CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.**

***Index Terms*—component, formatting, style, styling, insert**

I. INTRODUCTION

Breast cancer is the most frequent type of cancer in women globally, even more common than lung cancer.[1]. The GLOBOCAN study states that 10 million fatalities and 19.3 million cases of cancer were reported in 2020. It was anticipated that there would be 2.3 million new instances of breast cancer in women, surpassing the number of cases of lung cancer. The average death rate from breast cancer in women is 6.9%, demonstrating the seriousness of the disease's impact on women. One of the most frequent cancers in women is breast cancer, which may significantly affect their life. [2] The International Agency for Research on Cancer (IARC) has reported a 66% rise in cancer-related mortality worldwide since 1960. As a result, almost one in eight women in the US alone are predicted to acquire invasive BC at some point in their lives.[3] The best way to diagnose breast cancer is by raising awareness, doing early screening and diagnosis, and constantly improving treatment along with medical imaging examinations. While digital mammography, ultrasound (US), magnetic resonance imaging (MRI), and infrared thermography are among the imaging modalities utilized for diagnosis, mammography imaging is typically advised. Mammography generates high-quality pictures that make the interior structure of the breast visible. Mammograms can reveal several signs of breast cancer. Architectural deformities, masses, and macrocalcifications (MCs) are a few of them. Architectural distortions are determined to be less important than masses and MCs.[4] Determining whether the cells are malignant is crucial, and several kinds of images—known as biopsy images—are utilized to do this. Images of breast cancer are used to categorize the cells as benign, malignant, or noncancerous.[2] In this work, the diagnosis of breast cancer is reviewed systematically. Using machine learning (ML) and deep learning (DL), a thorough overview of the methods for analyzing breast cancer was provided in several publications. The purpose of this study is to provide a simplified overview of the historical development of breast cancer diagnostics while also addressing current technological breakthroughs and widely used procedures. It primarily focuses on a technique that combines clinicopathological data with genetic imprint analysis and is based on mammography and breast histopathology images. We reviewed research that was carried out over a number of years in order to understand the latest advancements in breast cancer diagnosis. This article comprises research that used several imaging modalities to diagnose patients. It is very crucial to diagnose breast cancer accurately from images. Many studies have been carried out up to this point, but many obstacles still need to be overcome, such as Hard to interpret due to interference of other tumor suppressor genes. [2] Dataset imbalance, inaccuracy, size, and computing efficiency. [3,4,6] Models scalability issues such as considering only particular models. [5] Not developing accurate algorithms instead of only using old models.[16] Limiting the generalizability such as not considering other datasets except particular datasets, only exploring machine learning(ML) approaches and not exploring other potential techniques, etc. [19] We reviewed several articles on the diagnosis and categorization of breast cancer to address the issues mentioned above. In addition to discussing several methodological options for handling unbalanced and sparse data and improving model scalability, we also cover various datasets that are currently accessible for the diagnosis of breast cancer. In addition, we go over several methods for pre-processing our data, how to identify and pick features, how to analyze various ML and DL classification models, and image modalities.

A. RESEARCH QUESTIONS

We shall address the following study questions in light of the aforementioned conditions. Determining the overall goal and anticipated results of a study depends heavily on the selection of research questions. So, in order to fulfil our study's primary goal, we created the following research questions:

RQ1: How datasets were collected and an idea of how and What are the common datasets used for machine learning and deep learning-based breast cancer detection?

RQ2: What are the image modalities like mammography, ultrasound, magnetic resonance imaging (MRI), and digital breast tomosynthesis?

RQ3 : What are the Advantages and Limitations of the Specific Methods used in our reviewed papers?

RQ4 : Which Machine learning algorithms works better on what kind of image modalities

So, based on our questions, the main contributions of our article are as follows:

A review of a total of 40 papers from 2020-2023 from various well-known sources like Elsevier, IEEE, and Springer has been covered regarding breast cancer detection and classification. The difficulties in diagnosing breast cancer and the researchers' answers have been systematically reviewed and analyzed. Shortcoming identification regarding these papers provided the probable solution for the real-world experiments, which will be beneficial for future research for the researchers along with giving insight to the doctors for better diagnosis and treatments.

The main goal is to enhance breast cancer detection and diagnosis by using this systematic review based on data extraction, feature selection, imaging modalities, and numerous machine learning algorithms.

II. DATASET

A. Prominent Datasets in Breast Cancer Research

There are many well-established datasets already out there for breast cancer, and at the same time, datasets are being created every day in the world from real patients in cancer hospitals. Researchers and practitioners use these datasets to identify breast cancer by creating ML models. These datasets contain information related to breast cancer, which plays a vital role for researchers.

Several Datasets are essential for the diagnosis of breast cancer because they provide important information on the complexities of the condition. The Wisconsin Diagnostic Breast Cancer (WDBC) dataset, which was created especially for Fine Needle Aspiration (FNA) treatments. The Mammographic Image Analysis Society (MIAS) is another significant dataset that includes X-ray pictures and provides comprehensive imaging to assist in the diagnosis of breast cancer. They can find mammography data in the Digital Database for Screening Mammography (DDSM), which serves as a comprehensive resource.

Other datasets, such as the Wisconsin Breast Cancer Dataset (WBCD), address FNA patients, and data for a variety of breast cancer studies may be found in Excel format at the Breast Imaging Archive at the University of Chicago (BIACH) and Research Institute (RI). Because they offer crucial details on the intricacies of the disease, a variety of datasets are crucial for the diagnosis of breast cancer. The Wisconsin Diagnostic Breast Cancer (WDBC) dataset is especially interesting as it was developed specifically for FNA therapies. Another notable dataset is the Mammographic Image Analysis Society (MIAS), which offers complete imaging to aid in the detection of breast cancer and include X-ray images. The Digital Database for Screening Mammography (DDSM) is a comprehensive resource that offers a plethora of mammography data to researchers and practitioners.

Other datasets, such as the Wisconsin Breast Cancer Dataset (WBCD), address FNA patients, and data for a variety of breast cancer studies may be found in Excel format at the Breast Imaging Archive at the University of Chicago (BIACH) and Research Institute (RI).

B. Image Modalities for Breast Cancer Diagnosis

Mammogram: This modality leads with 41 articles, indicating its prevalent use and significance in breast cancer diagnosis and research.Ultrasound: The second most focused modality, with 30 articles, reflecting its importance in breast cancer screening, particularly for

TABLE I: Description of Datasets

No.	Obtained from	Dataset	Image Modalities	Total Images	Limitation
1	University of Wisconsin Hospitals, Madison	WDBC	FNAs	Not actual images, numerical features extracted from digitized images of FNAs	Missing clinical context, No image in the dataset
2	Mammographic Image Analysis Society	MIAS	X-ray	332 mammographic images	Small size, single modality, Missing clinical context
3	Massachusetts General Hospital from the University of South Florida	DDSM	Digital Mammography	2620 scanned film	Limited clinical info, not representative of all population, single modality, limited annotation
4	Wisconsin Breast Cancer Dataset	WBCD	FNAs	Not actual images, numerical features extracted from digitized images of FNAs	
5	Indo-American Cancer Hospital and Research Institute IT department	BIACH & RI	Excel format	No actual images	
6	Surveillance, Epidemiology, and End Results managed by National Cancer Institute	SEER	Population-based cancer registry	Not actual images, database includes clinical and demographic info related to cancer cases	Incomplete information, lack of clinical information
7	Breast Research Group, INESC Porto, Portugal	INBreast	Mammogram	410	
8	Aachen University, Germany	IRMA	Mammogram	355	
9	INEGI, FMUP-CHSJ-University of Porto, Portugal and CETA-CIEMAT, Spain	BCDR	Mammogram	BCDR-D01(134), BCDR-D02(405)	Clinical context, Data Size and Diversity
10	Pathological Anatomy and Cytopathology (P&D) Lab, Brazil, University of Chicago Medical Center, Private-Sun Yat-sen University, Nanhai Affiliated Hospital of Southern Medical University, Engineering Department of Cambridge University	BreakHis	Biopsy, Mammogram	607	
11	Baheya Hospital, Cairo Egypt, King Hussein Cancer Center (KHCC), Jordan Hospital (JH)	BUSI	Ultrasound, Mammogram	252, 792	Specific Research Goals, Annotation quality

younger women or those with dense breast tissue. MRI: Though less common with 15 articles, MRI is crucial for detailed imaging, especially in high-risk patients.Histopathology: With 25 articles, it shows significant usage in the field, likely due to its role in detailed tissue analysis for cancer diagnosis. Breast cancer is one of the most common cancers in women worldwide and requires early detection and accurate diagnosis for effective treatment. In our study, we saw that various image modalities play significant roles in breast cancer detection and classification. The following imaging techniques discussed below each have unique capabilities and applications that are used for breast cancer screening, diagnosis, and treatment planning. The primary goal of these image modalities is to provide detailed and precise visualizations of breast tissue, enabling the detection of abnormalities such as tumors, calcifications, and other changes that might indicate the presence of cancer. From the initial screening to the precise staging of the disease, these imaging techniques are integral to the patient's diagnostic journey and subsequent treatment. There are several techniques of image modalities, we have tried to give a short overview of each:

i. Mammography:
- Digital Mammography: Captures breast images in digital format. It's the most common and widely used technique for breast cancer screening.
- Digital Breast Tomosynthesis (3D Mammography): Offers a 3-dimensional image of the breast, improving the detection of lesions, especially in dense breast tissue.

ii. Ultrasound:
- Used as a supplementary tool to mammography, especially useful in evaluating breast abnormalities and distinguishing fluid-filled cysts from solid masses.
- Helpful in dense breast tissue where mammography might not be as effective.

iii. Magnetic Resonance Imaging (MRI):
- Provides detailed images of the breast and is particularly useful in high-risk patients, for assessing the extent of cancer, and for monitoring response to therapy.
- Often used in conjunction with other imaging modalities for comprehensive evaluation.

iv. Computed Tomography (CT):
- Less commonly used for initial breast cancer screening but may be employed in certain cases to evaluate advanced cancer and monitor treatment response.

v. Positron Emission Tomography (PET):
- Often combined with CT (PET/CT) to assess cancer spread (metastasis) and monitor treatment effectiveness.

vi. Thermography:
- Uses infrared technology to detect heat and blood flow in breast tissues. It is not a standard screening tool but can be used as an adjunctive tool in some cases.

vii. Optical Imaging:
- Includes techniques like Near-Infrared Spectroscopy (NIRS) and Photoacoustic Imaging, used in research and experimental settings for breast cancer detection.

viii. Elastography:
- A specialized form of ultrasound that measures tissue stiffness, helping in distinguishing benign from malignant lesions.

Now, for this paper, we only considered four major mostly used modal images. A study[] went through around 1000 papers. We can see from the following graph, mammography, the most prevalent and standard

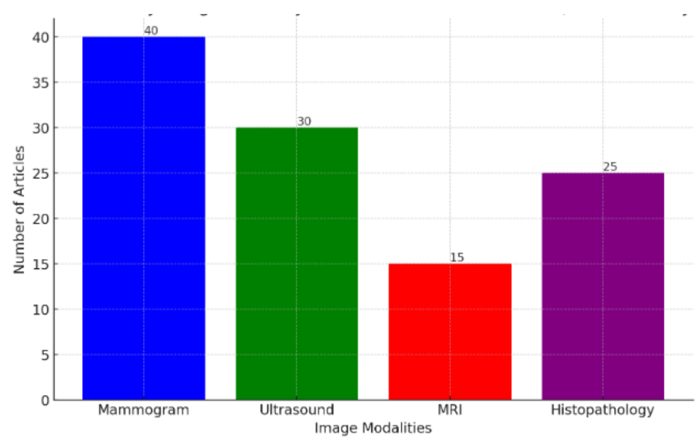

Fig. 1

modality, offers a fundamental approach to breast cancer screening. It's particularly effective in detecting early-stage cancers, even before they can be felt as lumps. Advances in this field, like digital mammography and digital breast tomosynthesis (3D mammography), have further enhanced its diagnostic accuracy, especially in dense breast tissues. From the graph, we can observe Mammogram: This modality leads with 412 articles, indicating its prevalent use and significance in breast cancer diagnosis and research. Ultrasound: The second most focused modality, with 305 articles, reflecting its importance in breast cancer screening, particularly for younger women or those with dense breast tissue. MRI: Though less common with 61 articles, MRI is crucial for detailed imaging, especially in high-risk patients. Histopathology: With 157 articles, it shows significant usage in the field, likely due to its role in detailed tissue analysis for cancer diagnosis.

III. MODELS USED IN CLASSIFICATION

A. MACHINE LEARNING MODELS

1) SVM: Several researchers approached support vector machines (SVM) for MRI breast cancer classification and diagnosis. SVM works best on small but complex datasets by finding the optimal hyperplane that separates different classes. It uses similar feature extraction like LR but it can incorporate techniques like PCA for dimensionality reduction to enhance efficiency. A study on breast cancer classification between natural and non natural using private dataset shows 89% accuracy using SVM. Another study utilized SVM to classify breast cancer on a private dataset with a higher accuracy of 94% . Some research employed SVM to detect benign and malignant classes using private dataset . On the other hand, another study applied subgroup classification using cross validation k-NN on a 200 private image dataset.

2) LOGISTIC REGRESSION: Logistic regression (LR) refers to the probability estimation of an event. LR acts as a detective for mammograms by looking at the pattern including textures and brightness. In order to predict the probability of malignant or benign class of breast cancer , LR built a linear model . LR makes best predictions on a small dataset achieving higher accuracy compared to other models. A study achieved accuracy of 85% on a dataset consisting 250 mammograms in classifying between benign and malignant classes. Another study achieved 82% accuracy on classification of several features extracted from mammograms.

3) DECISION TREE: Decision tree refers to a non-parametric supervised learning method that allows classification based on mammogram clues by using simple rules and taking its decision. DT shows some promising results in breast cancer classifications. A study employed a decision tree on 9 extracted features from ,mammogram to to classify breast cancer achieved an accuracy of 80% . Another study achieved an accuracy of 83% to classification breast cancer using the combination of decision tree and genetic algorithm .

4) NAIVE BAYS: Naive bayes classifier is also a supervised learning which estimates the chances of a tumour by observing all the clues present in a mammogram. Compared to other models it is quite fast but less used. A study using NV to classify breast cancer between benign and malignant classes on mammograms achieved an accuracy of 78% . Annonether study employed NV for classification of four different classes including normal, malignant mass, benign mass and indeterminate.

B. DEEP LEARNING MODELS

1) AlexNet: AlexNet is a classic Convolutional Neural Network which consists of several layers including Convolutional layer, max pooling and dense layer. It is considered a strong classification model because it works efficiently in complex and large datasets. The layers work as an artist's tools to identify the important textures and patterns from mammograms and accurately show the result in most cases.In Spite of the requires a large dataset and significant computational resources, AlexNet archives promising performance to classify breast cancer. A study achieved 96% accuracy to classify benign and malignant classes of mammograms using AlexNet . Similarly, Another study utilized AlexNet to classify breast cancer using ultrasound images achieved Specificity of 90% and sensitivity of 93%.

MODEL DISCUSSION(DL VERSION):

Ting, Tan, Sim [6] proposed a method using MIAS dataset for breast cancer classification. CNN improvement for breast Cancer Classification (CNNI-BCC), their suggested model has achieved remarkable result. Their research had 221 data's which produced great results: 89.47% sensitivity, 90.50% accuracy, 0.901 (± 0.0314) area under the curve (AUC), and 90.71% specificity. CNNI-BCC shows greater security, and dependability compared to other established models. CNN's ability to identify patterns in medical images improved the identification of possible breast cancer. Also, they address the issue of protecting patient data using great security. Although there may be issue like complications in the implementation, we need to recognize the effectiveness of CNNI-BCC. The scalability of model performance depends on the dataset quality and the amount, which is an ongoing process. This issue highlights the need to strike a careful balance in the continuous quest for more effective breast cancer detection between the promise of improved diagnostic accuracy and system complexity.

A.B.D.Zaher and Eldeib [7] used WBCD dataset in their research with a strong combination of backpropagation and deep belief network (DBN). While conjugate gradient learning achieved an excellent accuracy of 99.59% when analyzing 63.84% of the dataset, Levenberg-Marquardt learning produced an accuracy of 99.68% with the check on 54.9% of the cases. These results show the robustness of their model while identifying breast cancer. Moreover, their system's security improved by the combination of DBN and backpropagation which a result offered a strong defense to threats to patient's medical history.

Abunasser et al. [8] performed their research using following CNN models: Xception, InceptionV3, ResNet50, VGG16, MobileNet and their proposed model BCNN. Their suggested model achieved an outstanding 98.28% in the F1-score accuracy which clarify in identifying the breast cancer in eight different classifications. Models accuracy which then enhanced by the addition of a Generative Adversarial network (GAN). The model is scalable and flexible because it includes four image mag-

TABLE II: Summary of Breast Cancer Classification Studies

Paper Title	Model	Accuracy	Findings
A Framework for Breast Cancer Classification using Multi-DCNNs	DCNN	99.42%	CAD system assists diagnosis
Breast Cancer Classification Using Machine Learning	NB, KNN	96.19%, 97.51%	Dataset segmented to nine key criteria
A Comprehensive Review on Breast Cancer Detection, Classification and Segmentation Using Deep Learning	CLSTM	AC: 95.18%, PR: 95.44%, F1: 95.2%, RE: 95.2%	NA
Breast Cancer Classification Using Deep Belief Networks		Conjugate Gradient, Levenberg-Marquardt	99.59%, 99.68%
Convolution Neural Network for Breast Cancer Detection and Classification Using Deep Learning	Xception, InceptionV3, ResNet50, VGG16, MobileNet, BCCNN	97.54%, 95.33%, 98.14%, 97.67%, 93.98%, 98.28%	GAN for dataset
Classification and diagnostic prediction of breast cancer metastasis on clinical data using machine learning algorithms	DT	83%	accuracy dramatically improved by removing outliers from the blood profile data
Analysis of Breast Cancer Prediction Using Multiple Machine Learning Methodologies	RF	F1: 93.5%, Precision: 93.5%, Recall: 93.6%, Accuracy: 93.5%	RF, KNN outperform other DL methods
A Hybrid Machine learning Model for Timely Prediction of Breast Cancer	Hybrid ML	99.65%	enhance early detection and diagnosis
An Integrative Machine Learning Framework for Classifying SEER Breast Cancer	DT	98%	highlighted Feature selection importance
Breast cancer detection using machine learning approaches: a comparative study	SVM	99.7%	SVM achieved 97.7%; FN rate: 0.029, FP rate: 0.019
Comparison of machine learning models for breast cancer diagnosis	Gradient boosting	97.36%	LDA, LR, PA, NB, SVC achieve perfect precision
DL-based Automatic Diagnosis of Breast Cancer on MRI Using Mask R-CNN for Detection Followed by ResNet50 for Classification	R-CNN	96% sensitivity	Mask R-CNN for detection; ResNet50 for classification
Efficient breast cancer mammograms diagnosis using three deep neural networks and term variance	M-SVM	98% (80% training)	CNNs + TV outperformed ,Normal class max performance;

nification form 40k to 400k. Because of that, it is more complex to build. However, this model stads out from other proposed model because of its accuracy offering better security in the field of brest cancer classification and detection.

Sharma et al.[12] used WBCD dataset in their research and proposed a methodology involving several steps to identify breast cancer using a snapshot ensemble deep learning model and the dimensionality reduction technique which is t-distributed stochastic neighbor embedding (t-SNE). Which in result achieved 86.6%. They first preprocess the data and then to reduce the dimensionality they used the t-SNE which in result outperformed traditional algorithms. The snapshot ensemble achieved 81.2% accuracy, 91.1% precision, and 21.3% sensitivity, respectively. The proposed model ability to learn the local minima improved the snapshot ensemble model in the multiple training epochs.

Zhang et al. [16] used two datasets in their research and proposed a methodology to uses MRI images to identify breast cancer. There proposed model used Mask R-CNN to detect lesions, and ResNet50 is used to categorization. Which produced great results which includes 96% sensitivity in lesions detection. Also, Mask R-CNN accurately identified 101 out of 103 cancers. ResNet50 performed exceptionally well in the same data by an accuray of 99 out of 101 in which the model discovered malignancies. The author mentioned in their paper about the drawback of limitation of data. They also acknowledge the need of different datasets to compare their study. Also False positives detected by R-CNN were found to be from vessels, which can be identified using morphological operations. These limitations draw attention to the possible costs associated with putting such sophisticated systems into place, highlighting the necessity of careful dataset selection and continuous improvement to guarantee breast cancer diagnosis security and convenience.

Elkorany and Elsharkawy [17] used MIAS dataset and three CNN models InceptioV3, ResNet50, and AlexNet in their paper. CNN models performed as feature extractor in their study for mammography-based methodology. Their approach achieved classification accuracy (CA), averaging 97.81% for 70% of training, 98% for 80%, and ideal CA for 90% of training. Their suggested approach outperform other well established models. Specially, their models had perfect recall, specificity, precision, F1-score, and Area Under the Curve (AUC) in breast tissue types F and G. However, it's critical to examine the trade-offs between performance and simplicity of integration into current healthcare systems, taking into account the potential complexity and overhead in developing such advanced CNN models. The study focus on efficiency as well as dependability in breast cancer diagnostic systems.

Howard et al. [18] developed a deep learning model for predicting breast cancer recurrence test and risk where two nonconsecutive modules applied to image tiles extracted from digital slides. In where first module predicts the tumor likelihood and the second one predicts the recurrence scores. The combined DL model outperformed random chance in every case, as evidenced by the robust area under the precision-recall curve (AUPRC). The research notes that the lack of clinical-grade recurrence test findings accessible on TCGA may have an impact on how well the DL model performs when trained and validated using clinical-grade assay data. Furthermore, the validation dataset's limited number of recurrent episodes highlights the necessity for higher sample numbers to show clinical usefulness decisively. The dataset was created from patients coming from chemotherapy which may in result the need of modification to improve security.

Taheri and Golrizkhatami [20] used the BreakHis dataset in their research in where the proposed a new approach for diagnosing breast cancer. This dataset includes images of 7909 microscopic malignant breast tumors. Their proposed model based on pre-trained DenseNet201 architecture that exhibits improved dependability. One categorization system that is unique to magnification and another which is independent of magnification shows dependability in various diagnostic requirements. Their results outperform already established methods. The comprehensive evaluation of accuracy at the patient and image levels emphases the reliability of the proposed approach even more. The efficiency of these complex systems, however, highlights the narrow line between usefulness and seamless integration into the existing healthcare infrastructure and raises concerns about implementation challenges and potential costs. Despite the difficulties, authors represents a great advancement in the identification of breast cancer.

In comparison to existing approaches, the suggested method, proposed by Lee et al. [21], demonstrated better classification performance for breast cancer detection on DBT pictures. It does this by using a transformer endowed with split space-time attention to learn relations between neighbouring sections. Three criteria were used in the study to assess the classification performance: specificity at a set sensitivity, sensitivity at a fixed specificity, and the area under the receiver operating characteristic curve (AUC). 655 test trials comprising 163 cancer, 328 benign, and 164 normal patients made up the assessment dataset. The AUCs of several methods were compared using the DeLong test, and confidence limits were produced using an asymptotic normal approximation.

IV. RESULT AND ANALYSIS

A. *MACHINE LEARNING MODELS*

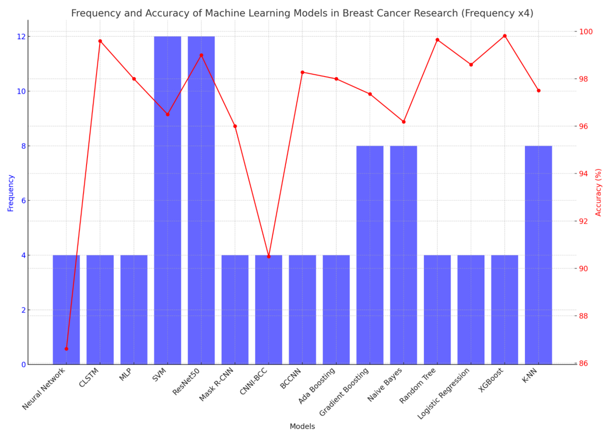

Fig. 2

The following chart shows that while some models are mentioned more frequently, this does not necessarily correlate with higher accuracy. For instance, the Support Vector Machine (SVM) and ResNet50 models are referenced twelve times each, indicating significant usage in the field. However, in terms of accuracy, the Convolutional Long Short-Term Memory (CLSTM) and the XGBoost models stand out with accuracies of 99.6% and 99.82% respectively, despite being mentioned less frequently. Interestingly, the Neural Network, a commonly used model in many fields, shows a moderate frequency of usage with a relatively lower accuracy of 86.6%. This could suggest that while Neural Networks are versatile, they might not always be the most accurate option for breast cancer research applications, or if we could have done further research been able to figure out on what conditions it will suit. On the other hand, models like the Gradient Boosting and Naive Bayes, which are mentioned 8 times, exhibit high accuracies of 97.36% and 96.19% respectively. This indicates their effectiveness in the specific context of breast cancer research, despite not being the most frequently used models. Overall, this analysis underscores the importance of considering both the frequency of use and accuracy when selecting machine learning models for research purposes. We can also say Random tree, which is a mixture of random forest and decision tree, since they both work in similar ways, have been given together, and have overall performed well, despite the fact that they are mostly used on word data rather than image data. It also highlights the diverse range of models being employed in the field of breast cancer research, each with its own strengths and limitations.

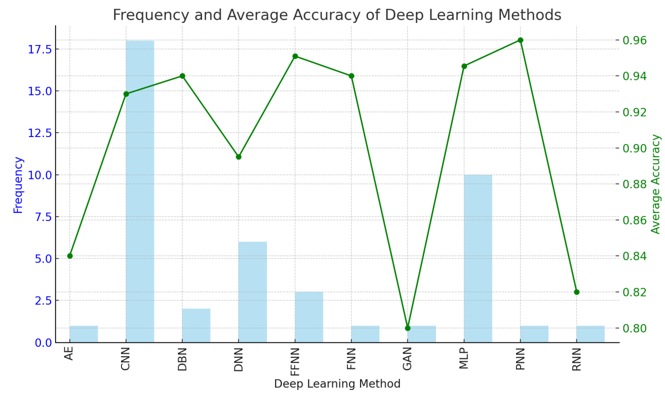

Fig. 3

B. *DEEP LEARNING MODELS*

A total of 40 papers were analyzed, The graph presents a dual perspective on the performance and popularity of various deep learning models used in image processing. It combines a bar chart, which indicates the frequency of use of each model, and a line graph that shows their average accuracy. The models considered are Feedforward Neural Network (FFNN), Convolutional Neural Network (CNN), Deep Neural Network (DNN), and Multilayer Perceptron (MLP). Feedforward Neural Network (FFNN): This model appears moderately frequently and demonstrates high accuracy. FFNNs are the simplest type of artificial neural networks. They consist of input, hidden, and output layers, where connections between nodes do not form cycles. In image processing, FFNNs can be limited due to their lack of spatial hierarchy, meaning they don't consider the spatial relationships between pixels. However, their simplicity makes them a good choice for less complex image datasets where contextual information is less critical. Convolutional Neural Network (CNN): CNNs are highly represented and show impressive accuracy. They are specifically designed for processing data with a grid-like topology, such as images. CNNs use convolutional layers, pooling layers, and fully connected layers to process and classify image data. Their ability to capture spatial hierarchies by learning from local patterns (like edges in early layers and more complex shapes in deeper layers) makes them exceptionally suited for image classification, object detection, and similar tasks. Their high accuracy and frequent use in the graph reflect their dominance in image processing tasks. Deep Neural Network (DNN): DNNs are used with a frequency similar to FFNNs but show a slightly lower average accuracy. The term 'DNN' can be somewhat ambiguous and often refers to any neural network with multiple hidden layers that can learn deep representations of data. In image processing, the depth of these networks allows them to learn complex patterns but may also make them prone to overfitting, especially when the amount of training data is limited. Multilayer Perceptron (MLP): MLPs are the least used among the four models and have the lowest average accuracy. An MLP is a class of feedforward neural network, consisting of at least three layers of nodes: an input layer, a hidden layer, and an output layer. Unlike CNNs, MLPs do not take into account the spatial structure in images, treating input pixels as flat vectors. This approach can be less effective for complex image tasks, where spatial relationships are crucial. The performance of these models can be attributed to their architectural characteristics and suitability for image datasets:

C. *IMAGE EXTRACTION MODELS ACROSS DIFFERENT IMAGING MODALITIES*

For this part, we have collected information on around 40 papers. Suggesting any particular model might be complex and depended on the type of image modalities we are using.In the following graphs, we included only the three main architectures that both had better performance and frequency in the papers. In the following sections we gave a review and analyzed of why the performed well

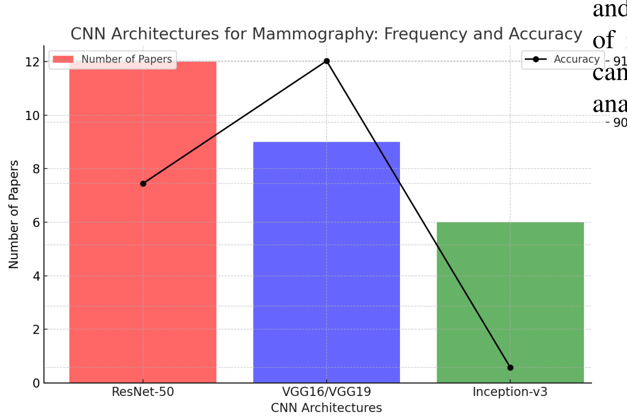

Fig. 4

1) MAMMOGRAPHY: Mammography images are high-resolution X-rays specifically designed to reveal the internal structure of breast tissues, which is essential for identifying abnormalities such as microcalcifications and tumors. Since these features are quite subtle and fine, they require a CNN architecture that can capture fine-grained details and learn from the high-contrast regions of the images. VGG16/VGG19 with their consecutive convolutional layers, are particularly adept at capturing the intricate spatial hierarchies present in mammography images. VGG networks are deep enough to learn complex patterns but have a uniform architecture that makes them easier to analyze and modify if needed. Their robustness has been proven in various image recognition tasks, and their ability to detect microcalcifications and delineate masses in mammograms has been demonstrated in medical imaging studies.

2) ULTRASOUND: Ultrasound imaging captures real-time images of the body's internal structures using sound waves, which are then reflected back to produce an image. These images tend to have a characteristic speckle noise and can vary significantly due to operator dependency and the settings of the ultrasound machine. U-Net: Originally designed for biomedical image segmentation, the U-Net architecture is particularly suitable for ultrasound images because of its symmetric expanding path that enables precise localization. This is important for ultrasound image analysis where the boundaries of organs and lesions may be blurred by speckle noise. U-Net's architecture can effectively learn from the limited data typically available in medical settings, making it ideal for distinguishing between benign, malignant, and normal ultrasound findings. MobileNet: In scenarios where computational resources are a constraint or real-time analysis is required, such as in portable ultrasound devices, MobileNet provides an efficient alternative. Its lightweight architecture, based on depthwise separable convolutions, is designed for mobile and edge devices, enabling quick and efficient processing of images without a significant drop in accuracy. This can be particularly beneficial in point-of-care ultrasound analysis, where rapid decisions may be required.

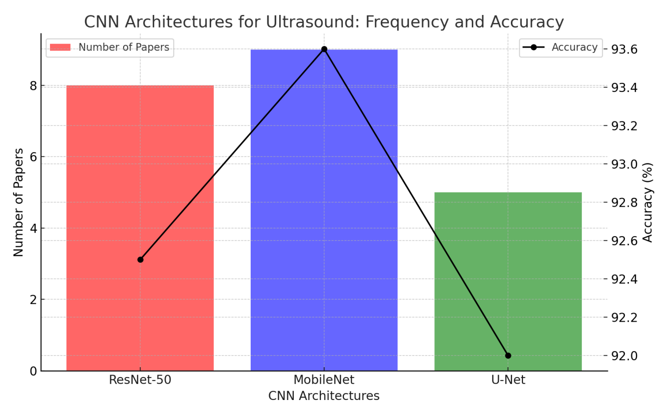

Fig. 5

3) MRI: MRI scans are complex and provide detailed images of the body's internal structures using magnetic fields and radio waves. They can produce a series of images that represent slices through the body, often resulting in a 3D dataset. The multiparametric nature of MRI, with different sequences like T1-weighted, T2-weighted, and FLAIR, highlights different tissue properties and pathologies. 3D CNNs: Given the volumetric

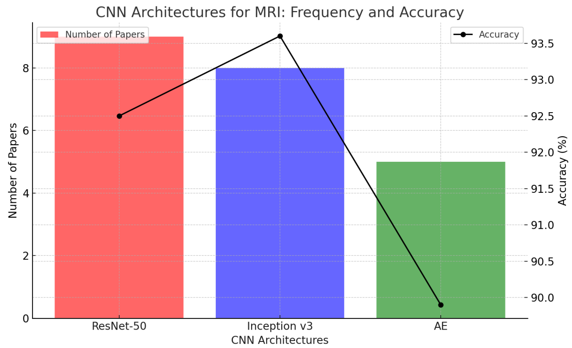

Fig. 6

nature of MRI data, 3D CNNs are capable of learning spatial hierarchies and patterns across consecutive slices, which is crucial for capturing the 3D structures and abnormalities present in the body. This capability makes 3D CNNs well-suited for MRI image analysis, as they can exploit the full spatial context of the data, potentially improving the classification of abnormalities in terms of their shape, size, and texture across multiple slices. Inception-v3/Inception-ResNet: These architectures are composed of modules that perform convolutions of different sizes concurrently, allowing the model to capture information at various scales and depths. This is particularly beneficial for MRI images, where the relevance of features can vary greatly depending on the sequence and the type of tissue or pathology. Inception networks can adapt to the diverse nature of MRI data, providing a flexible and powerful approach to capturing the most discriminative features for classification. In each case, the selected CNN architectures are chosen for their ability to learn complex patterns specific to the imaging modality, their proven track record in image classification tasks, and their adaptability to the requirements of medical image analysis.

V. LIMITATIONS

Many of the studies we went through relied on limited datasets, which may not represent the full spectrum of breast cancer cases. The lack of diversity in the data, especially in terms of ethnic and demographic representation, can limit the generalizability of the findings. Variability in image quality and lack of standardization across different imaging modalities can affect the accuracy of model predictions. This means that the pictures used for breast cancer studies, like X-rays or MRI scans, can look different depending on where and how they're taken. This variation can make it hard for computer models to correctly identify cancer because they might get confused by these differences. Most of the common dataset that we worked on was pre-processed. Though few papers around 20% used their own collected datasets.In simpler terms, for the best results in breast cancer research, it's important to have consistent and high-quality images and to know more about the patients' health background. Also we considered only image classification, though we narrowed it down for better coherence and kept it simple.

VI. FUTURE WORK

Our research was restricted to only English language related papers. There are papers on other languages also other geographical regions need to be taken into consideration for better and fair evaluation. Even though it is hard to suggest a particular model, overall, we can say Resnet 50 worked well on the 3 major image modalities. In future further work can be done to improve sector wise models to make it more scalable and real time. Moreover, we need to conduct our work on papers that have already done a real world implementation of their models, and we need to access the feasibility, scalability, and impact on patients. We also observed that there are few work done using the attention mechanism for image classification. Thus, this offers a chance for future researchers to take advantage of attention to raise the precision of deep learning models.

REFERENCES

[1] Anothaisintawee, T., Wiratkapun, C., Lerdsitthichai, P., Kasamesup, V., Wongwaisayawan, S., Srinakarin, J., Hirunpat, S., Woodtichartpreecha, P., Boonlikit, S., Teerawattananon, Y., and Thakkinstian, A. (2013, May 23)," Risk Factors of Breast Cancer. Asia Pacific Journal of Public Health," 25(5), 368–387. https://doi.org/10.1177/1010539513488795

[2] Mourouti, N., Kontogianni, M. D., Papavagelis, C., and Panagiotakos, D. B. (2014, September 8). "Diet and breast cancer: a systematic review. International Journal of Food Sciences and Nutrition", 66(1), 1–42. https://doi.org/10.3109/09637486.2014.950207

[3] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[4] Chaurasiya, S., and Rajak, R. (2023, June 21). "Comparative Analysis of Machine Learning Algorithms in Breast Cancer Classification. Wireless Personal Communications ", 131(2), 763–772. https://doi.org/10.1007/s11277-023-10438-9

[5] Li, T., Nickel, B., Ngo, P., McFadden, K., Brennan, M., Marinovich, M. L., and Houssami, N. (2023, February). "A systematic review of the impact of the COVID-19 pandemic on breast cancer screening and diagnosis". The Breast, 67, 78–88. https://doi.org/10.1016/j.breast.2023.01.001

[6] Ye, F., Dewanjee, S., Li, Y., Jha, N. K., Chen, Z. S., Kumar, A., Vishakha, Behl, T., Jha, S. K., and Tang, H. (2023, July 6). "Advancements in clinical aspects of targeted therapy and immunotherapy in breast cancer ". Molecular Cancer, 22(1). https://doi.org/10.1186/s12943-023-01805-y

[7] El Haji, H., Souadka, A., Patel, B. N., Sbihi, N., Ramasamy, G., Patel, B. K., Ghogho, M., and Banerjee, I. (2023, August). "Evolution of Breast Cancer Recurrence Risk Prediction: A Systematic Review of Statistical and Machine Learning–Based Models". JCO Clinical Cancer Informatics, 7. https://doi.org/10.1200/cci.23.00049

[8] Nasser, M., and Yusof, U. K. (2023, January 3). "Deep Learning Based Methods for Breast Cancer Diagnosis: A Systematic Review and Future Direction". Diagnostics, 13(1), 161. https://doi.org/10.3390/diagnostics13010161

[9] Zheng, D., He, X., and Jing, J. (2023, January 4). "Overview of Artificial Intelligence in Breast Cancer Medical Imaging ". Journal of Clinical Medicine, 12(2), 419. https://doi.org/10.3390/jcm12020419

[10] Nemade, V., Pathak, S., and Dubey, A. K. (2022, April 11). "A Systematic Literature Review of Breast Cancer Diagnosis Using Machine Intelligence Techniques". Archives of Computational Methods in Engineering, 29(6), 4401–4430. https://doi.org/10.1007/s11831-022-09738-3

[11] Khan, N., Adam, R., Huang, P., Maldjian, T., and Duong, T. Q. (2022, November 21). "Deep Learning Prediction of Pathologic Complete Response in Breast Cancer Using MRI and Other Clinical Data: A Systematic Review". Tomography, 8(6), 2784–2795. https://doi.org/10.3390/tomography8060232

[12] Loizidou, K., Elia, R., and Pitris, C. (2023, February). "Computer-aided breast cancer detection and classification in mammography: A comprehensive review ". Computers in Biology and Medicine, 153, 106554. https://doi.org/10.1016/j.compbiomed.2023.106554

[13] Amethiya, Y., Pipariya, P., Patel, S., and Shah, M. (2022, May). "Comparative analysis of breast cancer detection using machine learning and biosensors" . Intelligent Medicine, 2(2), 69–81. https://doi.org/10.1016/j.imed.2021.08.004

[14] Arun Kumar, S., and Sasikala, S. (2023, January). "Review on Deep Learning-Based CAD Systems for Breast Cancer Diagnosis". Technology in Cancer Research and Treatment, 22, 153303382311779. https://doi.org/10.1177/15330338231177977