A Comparative Study Of Machine Learning And Deep Learning Models For Breast Cancer Detection

MSc Research Project

MSc Data Analytics

Anjaly Antony Achandy

Student ID: X21223661

School of Computing

National College of Ireland

Supervisor: Musfira Jilani

National College of Ireland

MSc Project Submission Sheet

School of Computing

Name:	Anjaly Antony Achandy		
Student ID:	X21223661		
Programme:	MSc Data Analytics	**Year:**	2024
Module:	MSc Research Project		
Supervisor:	Musfira Jilani		
Submission Due Date:	12/12/2024		
Project Title:	A Comparative Study Of Machine Learning And Deep Learning Models For Breast Cancer Detection		
Word Count:	5825	**Page Count:**	23

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Anjaly Antony Achandy
Date: 12th December 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	✓
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	✓
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	✓

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

A COMPARATIVE STUDY OF MACHINE LEARNING AND DEEP LEARNING MODELS FOR BREAST CANCER DETECTION

Anjaly Antony Achandy

21223661

ABSTRACT

Breast cancer affects women worldwide and is still a leading killer disease. This is why early diagnosis is extremely important to try and raise the percentage of people surviving this disease. Mammography is the main imaging method used to detect breast cancer, although its interpretation is difficult because breast tissue density varies, which results in missed detection or callback. Recently, the methods of deep learning, especially convolutional neural networks (CNNs), present in most imaging tasks, offer the potential to improve the accuracy rate and decrease the workload for radiologists. The particular dataset for the current work is the MIAS (Mammographic Image Analysis Society) dataset accessible on Kaggle including 322 mammograms labeled with classes (normal, benign, malignant), types of tissue density, and the lesion's coordinates.

1 INTRODUCTION

It is important to know the symptoms and diagnose the illness at a very early stage, and especially it is more manageable if the cancer type is diagnosed early. Mammography is currently the gold standard in breast cancer screening, while the interpretation of mammograms is difficult, more so when working with a large number of scans. Inaccuracies linked to fatigue and actual human mistakes may result in overlooking and misinterpretation of various manifestations (Marks *et al.*, 2000).

CAD systems are now integrated into almost every new mammography system to help locate and categorize abnormalities in digital mammograms. Even now, often, it remains difficult to determine whether the tumor is benign or malignant, although each one requires different methods of treatment. The existing ones for example tissue biopsy give a correct diagnosis but are expensive, invasive and commonly used which makes them less effective for population screening.

CAD systems and early screening by picture archiving and communications system can greatly diminish breast cancer mortality. Current approaches rely on traditional methods that involve hand–labeling and can be highly time–consuming in large datasets/ However, mammogram images analysis using current machine learning models such as Convolutional Neural Networks (CNNs) can accurately and efficiently solve this challenge (Schueller *et al.*, 2008). Such models assist radiologists in zoning the malignancy regions more accurately, thereby excluding other possibilities of true positive and hence unwarranted treatment.

CAD systems and advanced imaging technologies enhance the basic screening mammography and diagnostic accuracy; therefore, CAD systems and advanced imaging technologies provide better opportunity for accurate early detection of breast cancer and reduce mortality rates and increase patient survival.

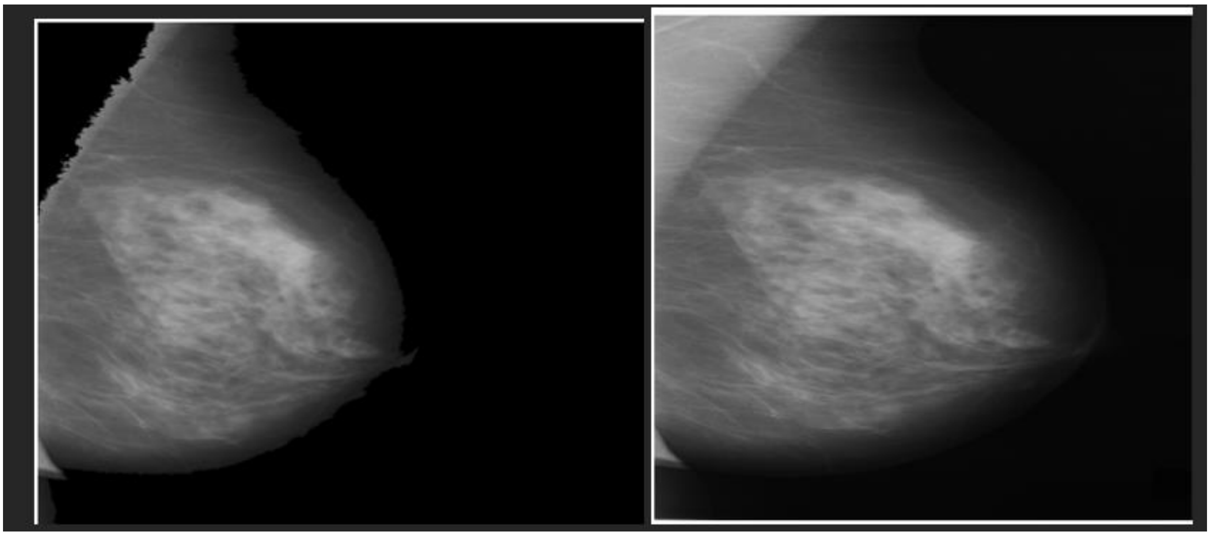


Figure 1: Brest Cancer Images (mammographic)

These conditions arising from the current state methods for breast cancer treatment open up a gap for early detection and achievement of a mass audience appeal for solving breast cancer. However, due to the computational capabilities of the computer vision and machine learning practices biomedical image detection has been a strong application of artificial intelligence. In addition, due to the highlighted obstacles, Medical practitioners can eliminate diagnostic differences with the help of artificial intelligence technology (Ferlay *et al.*, 2020). As pattern recognition and machine learning have emerged to classify pictures of breast cancer histology, a number of research using manually created or artificially improved features have been developed. The section shows that numerous initiatives have failed to meet the requirements necessary for building dependable automated learning systems; however, numerous technological enhancements accompanied by computational solutions that make use of deep and machine learning have been successfully integrated. The authors of the articles include Zebari D. et al. (2020) aimed at the creation of an autonomous segmentation pipeline through deep learning. However, they faced some challenges; regarding the generalization of the results to real-world data and low conversion accuracy that made their pipeline unduly for usage in the research labs.

However, convolutional neural networks (CNNs) for medical image processing have had quite a few achievements in the field. The CNNs have come in handy when it comes to duplicity of filter mappings in the identification of tumours and their invasiveness since researchers have been able to draw very clear distinctions in this field. CNNs have been combined with machine learning and transfer learning approaches including Support Vector Machines as used by Sarosa et al. (2018), while integration of KNN, Random Forest, and Decision Tree as proposed by Liu et al. (2019). Likewise, both of these works appeared in print and within scholarly academic journals. Furthermore, Sabre et al. (2021) proposed a transfer learning model with its architecture based on dense convolutional neural network.

The general objective of this research is to create a deep learning breast cancer detection and diagnosis model using mammography images in MIAS database. This paper focuses on the assessing an AI framework for preliminary test to diagnosis breast cancer using data from Kaggle's MIAS dataset. The MIAS dataset has 322 MLO mammography images each has expert labelled with regard to extent of breast abnormality. Images range from normal glandular tissue as well as benign or malignant breast lesion. MIAS dataset contains left and right breast mammographic images for 161 patients with well known pathology. The methodology under development is such that, after successful training and validation, the CNN will act as the classifier of normal and abnormal cases in the MIAS digitized mammography. More experiments also will help in distinguishing between benign and malignant anomalies will also be observed. Therefore, regarding the model's general performance as well as the efficiency and performance parameters of their best model of accuracy, sensitivity, specificity, and AUC, certain tests will be conducted. Once the training and testing to the everyone's satisfaction has been done, the CNN model will provide an AI instant triage of cases that need further tests or biopsy (Omonigho *et al.*, 2020).

As, at this stage of the development of the technology most of the mammography CAD systems reveal comparably high false positive rates, the issue of specificity, which has been chosen for this project, will be discussed with some priority at the cost of moderate decrease in sensitivity. This will also enable the interpretation of the CNN model by identifying the imaging features in which the model relies on for classification. Thus, the proposed CAD system should improve the performance of both the radiologist as well as the system and should increase the confidence of the patients as well as the radiologists regarding the use of AI in the screening of breast cancer (Singh, 2020). The project will try and achieve detection accuracies that are similar to those achieved in the case of the radiologist, while the actual inputs given to the AI would be actual mammography images.

The study aims to achieve the below-enlisted novelties:

a. Combines features extracted through multi-layer transfer learning pipelines namely VGG16 with processed neural backbone.
b. Merges feature extraction and selection into a hybrid neural network architecture featuring Deep CNN and Attention backbones forming the RCANN or Residual Convolutional Attention Neural Network.
c. Leverages state-of-the-art datasets such as MIAS for comprehensive analysis.
d. Conducts evaluations using metrics including accuracy, precision, recall, and F1 score.
e. Incorporates statistical distribution tests to validate the model's performance robustness.

2 LITERATURE REVIEW

The early detection of breast cancer poses a real challenge to the detection and reduction of fatality of breast malignant tumors. This remains the leading health challenge for the detection and identification of cancer tumor cell regenerative abilities. The computational methods have been illustrative in the detection procedure however the absence of a residual attention mechanism creates an inability towards explainable AI solutions.

2.1 Traditional Computer-Based Approaches for Breast Cancer Detection

First, main traditional machine learning techniques have been extensively used in the detection pipeline to provide a realistic reference point. According to Sarosa et al. (2018) proposed a feature extraction using the Gray Level Co-occurrence Matrix (GLCM) kernel simulation in conjunction with a feature-discriminant model, a support vector machine (SVM). The SVM forms a biased kernel that maps areas of uncertainty, differentiate between the benign samples that resembles the malignant ones. Any samples that look like they are more like these are placed along the bias of the discriminitive boundary of the kernel (Sarosa *et al.*, 2018).

Low-income women across the United States need better early detection protocols for mammary and cervical cancer, According to Marks *et al.*, (2000). Their plan includes four key interventions: enhancement of program reach, structural changes for patients, mobile clinics and additional clinic hours, screening rebates coupled with quality service provides. Included are best practices that offer clear policies on how this can be done and how different organizations should engage the communities to deliver services that are culturally sensitive to the requirements of minority classes of the population. All of these strategies were developed to reach the women who have had limited access to screening services to help reduce health disparities (Mark *et al.*, 2000). Additional studies of these interventions might help in determining how to focus future cancer prevention and control for high-risk populations.

Schueller and coworkers assess the accuracy of large-core needle biopsy in achieving correct breast lesion diagnosis, the estimation of the average number of samples per case, and correspondence with surgery histology. The authors performed a chart review of ultrasound-guided biopsies of 167 breast lesions in which a 14-gauge needle and surgical excision were done. For malignancies; the sensitivity of biopsy was 98% while Specificity was 67%; the accuracy of biopsy was 90%. They infer that while the positive correlation of large-core biopsy with final surgical pathology is high, the sampling error rates are high when trying to upgrade invasive carcinoma components, or DCIS after operation. Some of the weaknesses include their limited evaluation of the various systems of rewards across other than one organization.

2.2 Machine Learning Approaches for Breast Cancer Detection

The range of application of machine learning models has been widely applied in different fields. Amrane et al., (2018) applied breast cancer classification using machine learning and employed basic feature extraction methods. Later, Wang et al. (2019) and Khan et al. (2019) extended this to enhance MIAS mammographic samples using discriminative kernels to combine features. This method allows for capturing a high number of features thereby enriching an understanding of the results, as well as improving the creation of decision boundaries from the dense features. Therefore, the ratio of erroneous output samples to the number of correct samples increased by 13% compared to the traditional approach.

Ferlay et al., have put forward the estimates of cancer incidence and mortality rates for the year 2020, according to the data available with IARC. The report gives indicative new cancer incident rates by country, cancer type, gender, and age, and other highlights include the incidence drivers such as; population growth, and aging especially in low-income countries. It is an essential reference source for investigators planning and implementing cancer prevention and treatment programs. It provides an opportunity to monitor trends in cancer incidence and mortality in a particular region in real time.

Lei et al. (2021) discussed global breast cancer incidence and mortality trends throughout 2000-2020 using cancer registry data. They noted that breast cancer incidence had been increasing almost in all areas of the globe while mortality remained either declining or stagnating, which they attributed to improved awareness and management. However, the occurrence has increased significantly in transitional countries without a source. Nevertheless, the structure remained the same, and the differences were reflected in the degree of dispersion by geographical areas as well as the level of economic development. Above all, the study shows the rising trend in the incidence and mortality due to breast cancer not only in the developed countries but also in the developing countries hence, there is a need to avail early detection and quality treatment to women in the whole world. Subsequent studies are encouraged to identify potential reasons for these trends to guide Callan's resource management plan.

2.3 Deep Transfer Learning Approaches for Breast Cancer Detection

Deep architectures allow not only the computational analysis of the given multi-spectral mammographic images, in terms of their multi-layer kernels for their interpretation or sequence of decision boundaries. This is even made easier by transfer learning that allows the reuse of weights and bias of other networks from large datasets like ImageNet among others. Saber et al. (2021) and Singh et al. (2020) used AlexNet and DenseNet to extract features of images whereas imbalanced dataset problems were also considered.

Nevertheless, these pioneering models showed promise on the grounds of their capacity to quickly and efficiently analyze features, though, at the same time, they struggled with problems, such as overfitting and the lack of means that could assist in performing attention-based simulation to contribute to enhancing focus on important regions in images.

This utilizes the CNN approach under discussion and will be applied to breast cancer detection derived from mammography images. The authors use a pre-trained VGG 16 model and then choose among all features using recursive feature elimination. They have used 700 images in their dataset. They thus obtain an accuracy of 98.57%, they therefore better comparable models in terms of performance. Anyhow, more images particularly difficult ones may help to increase confidence in performance even more. Specifically, it is demonstrated that CNNs can learn subtle features useful for automatic breast cancer detection. Based on this line of thinking, CNN-based systems ought to be able to assist the radiologists in taking up their loads. As such, the number of sample images analyzed and more tuning of the models support the need to explore the use of CNN mammography analysis in improving the diagnosis of breast cancer.

In an article in IEEE Access, Jan 2019, Sun et al developed a method to diagnose mammographic images as normal, benign, or malignant using multi-view convolutional neural networks (CNNs). Their approach used both craniocaudal (CC) and mediolateral oblique (MLO) mammograms of the same breast and provided a better classification accuracy than a single view CNN. On the DDSM database, they compared several dual-input CNN architectures which all exceeded over 90% and were higher than other methods. They confirm that combining information from more than one view of a mammogram is very effective in the classification of breast lesions.

2.4 Literature Gap

The study finds numerous aspects while detecting breast cancer, specifically in perceiving correct and explainable AI solutions. On the other hand, conventional machine-learning methods are been used, which are mainly based on biased kernel-based models and lack of focus on improper datasets. The approaches of deep learning give high accuracy, insufficient attention, and encounter overlearning to emphasize key areas in mammographic images (Sun et al., 2019). There are some typical gaps in combining unused attention mechanisms with the frameworks of deep learning to improve both robustness and interpretability. Moreover, low-resources settings need more comprehensive protocols to ensure equal access to identify former detection and services regarding treatment.

2.5 Conclusion

The literature review highlights the potential work done across the domains of traditional cognitive approaches, machine learning approaches, deep–transfer learning solutions, and optimization-based approaches. The Research question which is thus a summarisation of the current state-of-the-art problems is defined below:

A. The Comparison between Machine Learning and Deep Learning: While algorithms such as SVM and logistic regression are popular for cancer detection, little research goes directly against more sophisticated modern deep learning models such as CNNs. As for the second approach, the question of how its performance might trade off with interpretability compared to the first approach is not well understood.
B. The Absence of Transformer-Based Methods: However, as seen in the prior section CNNs are widely used for the detection of breast cancer whose review with transformer models which have significantly changed natural language processing is lacking in these studies though the transformer models are recently emerging as promising models in image processing domain (Amrane et al., 2018). This seemingly constitutes an important research gap in the literature since the transformers may well provide better performance in image classification.
C. The Lack of Explainable AI (XAI): CNNs are mainly classified as black-box models and − to the best of our knowledge − effective XAI approaches for breast cancer detection are nearly non-existent. The clinicians need models with high accuracy, but models that also show how they arrived

at that conclusion in the first place to have trust in the models (Wang et al., 2019). This literature review confirms that deep learning techniques have enhanced breast cancer detection, but more research is required to close these gaps, including model interpretability as well as the novel architecture of transformers.

3 Methodology

3.1 Dataset Description

The research aims to formulate the use of mammographic imagery for breast cancer detection. The study thus uses the MIAS dataset or Mammographic Image Analysis Society. The dataset contains mammographic samples which are annotated by radiologists for breast cancer detection. The dataset comprises 1024 x 1024 images, which are classified into background tissue i.e. Fatty, Fatty glandular, and dense glandular; Severity classes Benign and Malignant classes.

The MIAS dataset consists of a total of 4680 images, will split the dataset into training, testing, and validation sets based on the specified proportions (65–72% for training, 15–20% for testing, and the remainder for validation) in a 10-fold cross-validation setup. Below is the table with actual image counts for each split in Table 1: Cross Validation - Train, Test, and Validation sets

Table 1: Cross Validation - Train, Test, and Validation sets

Fold	*Train (%)*	*Train (Samples)*	*Test (%)*	*Test (Samples)*	*Validation (%)*	*Validation (Samples)*
Fold 1	65	3042	20	936	15	702
Fold 2	66	3089	19	889	15	702
Fold 3	67	3136	18	842	15	702
Fold 4	68	3178	17	796	15	702
Fold 5	69	3225	16	749	15	702
Fold 6	70	3276	15	702	15	702
Fold 7	71	3323	15	702	14	655
Fold 8	72	3369	15	702	13	609
Fold 9	72	3369	16	749	12	562
Fold 10	72	3369	17	796	11	515

3.2 Machine learning

The ability of computational methods to draw discriminative kernel boundaries allows the model to predict benign and malign samplings in a much faster and easier way in comparison to the current tissue biopsy method. Machine learning is the early exploration of artificial kernel boundary formation from a trained data sampling method. In this study, it is aim to postulate a comparison between bagging and boosting methods. The paper extends comparison from the decision tree a traditional machine learning method where the tree formulates an entropy-based path system from various pixel parameters to evaluate the final classification. Further, utilize random forest, a cluster of decision tree structures. It is also an extension of bagging methods with a single base machine learning method. A multi-algorithm bagging setup is also utilized in the voting classifier method. Which uses soft aggregation for the cumulation of the output stimulus. Finally, the boosting method of adaptive boosting is utilized for the updating of enhanced boosting meta learner. Which uses log loss compounding to reduce the dynamic prediction fault of the algorithm.

3.3 Deep Learning

The deep learning impact of the extension of computer vision has been of great significance. As discussed thoroughly across the literature review the clear absence of comparative machine and deep learning paved the way for a comparative study on the MIAS dataset. To conduct the comparative analysis, the study aims to compare artificial neural networks, an ANN is a multi-layer perceptron with one hidden layer, which consists of linear neurons followed by input and output layers across both ends.

3.4 RCANN or Residual Convolutional Attention Neural Network

RCANN is a colligation of convolutional and attention neural network. The network is capable of addressing the vanishing and exploding gradient issues which were seen as a continual problem across machine and deep learning methods. The ability to integrate residual convolutional attention mechanisms. The model aims to achieve spatial attention for decoding features within the malignant samples which form the genesis of detection.

The network diagram consists of the VGG 16 module followed by CNN and Attention units as showcased. Figure 2: RCANN architecture

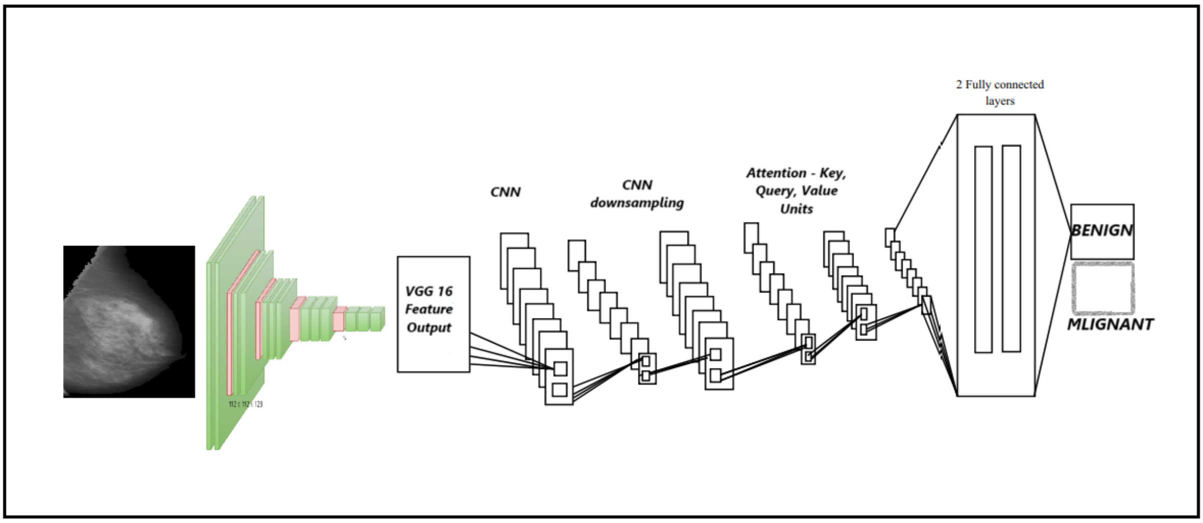


Figure 2: RCANN architecture

Equation i showcases the residual connection of the RCANN module followed by the Equation ii The SoftMax activation function with channel-wise attention network.

$$y = \mathcal{F}(x, \{W_i\}) + x$$

Equation i

$$z = y \cdot \operatorname{softmax}\big(A(y)\big)$$

Equation ii

Table 2: RCANN Architecture

Layer (Type)	*Output Vector*	*Parameters*	*Connection*
Input Layer (input_1)	(None, 100, 100, 3)	0	-

VGG16 Model (vgg16)	(None, 3, 3, 512)	1,47,14,688	input_1
Batch Normalization (batch_normalization_1)	(None, 3, 3, 512)	2,048	vgg16
Convolutional Layer 1 (conv2d_1)	(None, 3, 3, 64)	32,832	batch_normalization_1
Convolutional Layer 2 (conv2d_2)	(None, 3, 3, 16)	1,040	conv2d_1
Convolutional Layer 3 (conv2d_3)	(None, 3, 3, 1)	17	conv2d_2
Convolutional Layer 4 (conv2d_4)	(None, 3, 3, 512)	512	conv2d_3
Multiplication Layer (multiply_1)	(None, 3, 3, 512)	0	conv2d_4, batch_normalization_1
Global Average Pooling (global_average_pooling2d_1)	(None, 512)	0	multiply_1
Global Average Pooling (global_average_pooling2d_2)	(None, 512)	0	conv2d_4
Total Parameters	-	1,48,17,059	-
Trainable Parameters	-	1,00,835	-
Non-trainable Parameters	-	1,47,16,224	-

As showcased in Table 2: RCANN Architecture, it combines the CNN model – the VGG16 network and layers added for the extraction of rather refined spatial and non-kernel-based features. The model takes the images for input which are initially of size 100×100×3.

After that, the network employs the VGG16 model to parse the input and generate high-level features which is a 3x3x512 output. The next layer is the batch normalization layer which brings coherence to the measures and normalizes the activations.

The convolutional layer flattens the depth to 3x3x128, then another convolutional layer raises it back to 3x3x512 to improve the features. The global average pooling is carried out twice: first on the output of the multiplication layer and second on the last convolutional layer, and then results in a vector of 512 dimensions. The model's total weights are 14,817,059 this is made of 100,835 trainable weights and 14,716,224 that are nontrainable.

4 RESULTS AND DISCUSSION

This section evaluates the model's parameters across various subheadings such as machine learning, deep learning comparison, RCANN performance evaluation, explainable AI visualization, and finally state-of-the-art comparison. To achieve the same the study illustrates the use of various metrics showcased in Table 3: Metrics

Table 3: Metrics

Metric	***Formula***

A	$\frac{(T-)+(T+)}{(T-)+(T+)+(F+)+(F-)}$
P	$\frac{(T+)}{(T+)+(F+)}$
R	$\frac{(T+)}{(T+)+(F-)}$
F1S	$\frac{2*P*R}{P+R}$
SP	$\frac{(T-)}{(T-)+(F+)}$
Sensitivity	$\frac{(T+)}{(T+)+(F-)}$

The following Table 3: Metrics provide several performance measurements that are used in classification model evaluation.

4.1 Machine Learning Comparative Analysis

The Machine learning techniques that are defined in the methodology section are compared across both class-wise and generic comparison methods.

Table 4: Machine Learning Analysis

Metric	***Random Forest***	***Decision Tree***	***Voting Classifier***	***Adaptive Boosting***
ROC AUC Score	0.9391	0.9312	0.8639	0.5652
Cohen's Kappa	0.8782	0.8624	0.7278	0.1303
Jaccard Similarity	0.8852	0.8713	0.7603	0.3593
Precision (BBC)	0.9241	0.9342	0.8494	0.5391
Precision (MBC)	0.9553	0.9283	0.8796	0.6958
Recall (BBC)	0.9568	0.9278	0.8846	0.8987
Recall (MBC)	0.9214	0.9346	0.8432	0.2316
F1-Score (BBC)	0.9402	0.931	0.8667	0.6739
F1-Score (MBC)	0.938	0.9314	0.861	0.3475
Accuracy	0.9391	0.9312	0.8639	0.5652

As per Table 4: Machine Learning Analysis, model Random Forest has achieved a ROC AUC score of 0.9391, Cohen's Kappa of 0.8782, Jaccard Similarity of 0.8852, and an Accuracy of 0.9391. Model Decision Tree has achieved a ROC AUC score of 0.9312, Cohen's Kappa of 0.8624, Jaccard Similarity of 0.8713, and an Accuracy of 0.9312.

Model Voting Classifier has achieved an ROC AUC score of 0.8639, Cohen's Kappa of 0.7278, Jaccard Similarity of 0.7603, and an Accuracy of 0.8639. Model Adaptive Boosting has achieved an ROC AUC score of 0.5652, Cohen's Kappa of 0.1303, Jaccard Similarity of 0.3593, and an Accuracy of 0.5652.

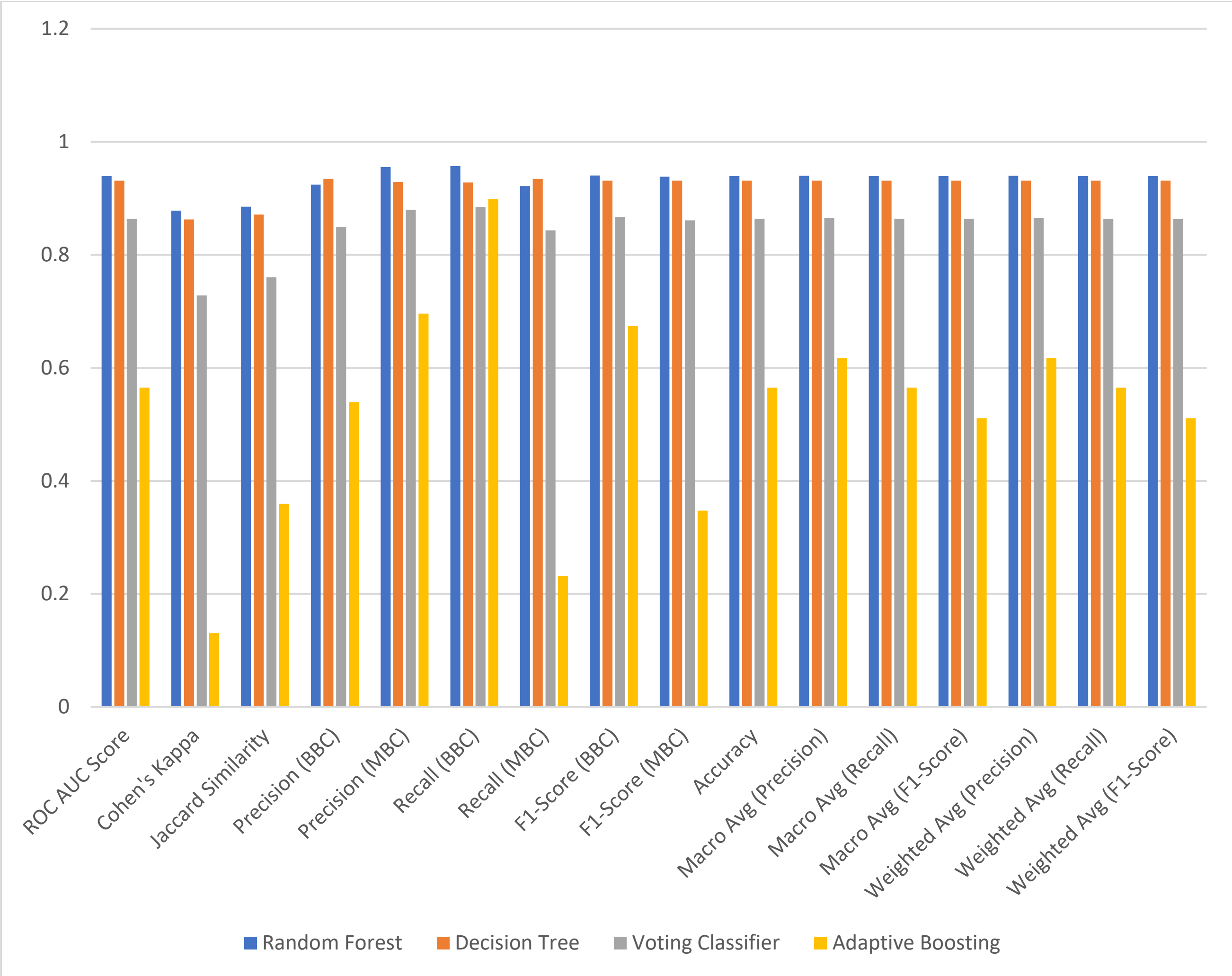


Figure 3: Machine Learning based comparison

The Figure 3: Machine Learning based comparison Describes the comparison of all algorithms across various comparative metrics. Allowing a view of global metrics and their performance.

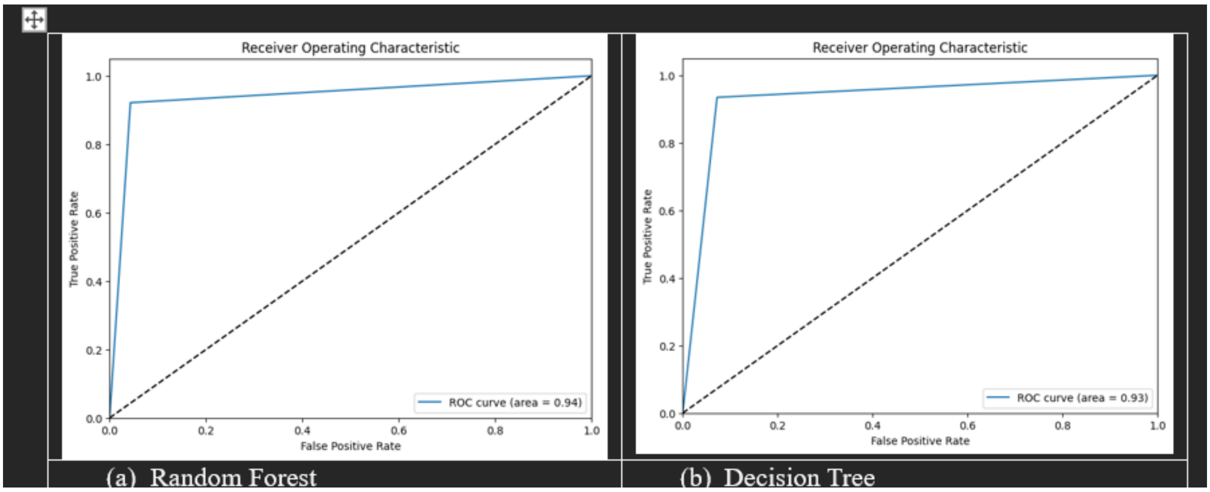


Figure 4: Receiver operating curve

Figure 4: Receiver operating curve Showcases the ROC curve and area under the curve for all the different machine learning algorithms. Figure 5: Confusion Matrix for machine learning models Showcases the Confusion matrix for all the different machine-learning algorithms.

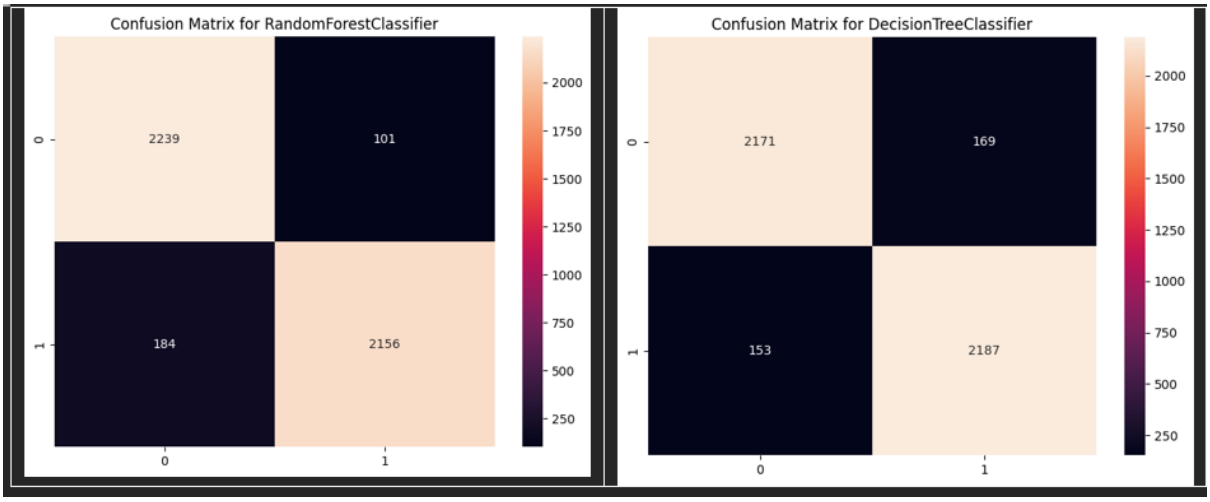
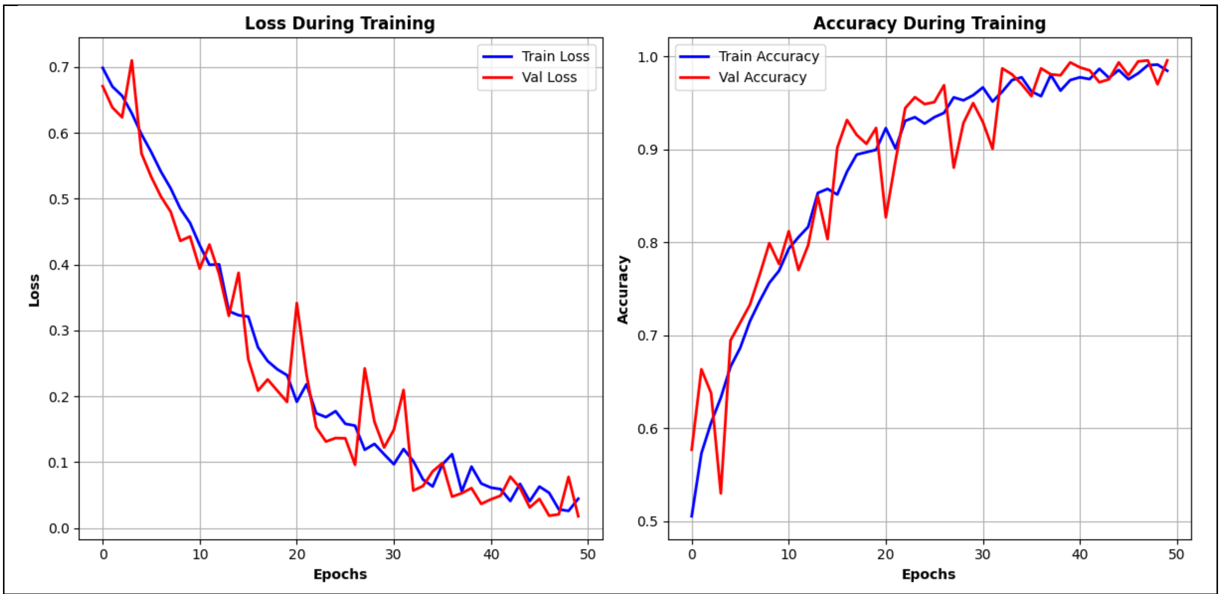


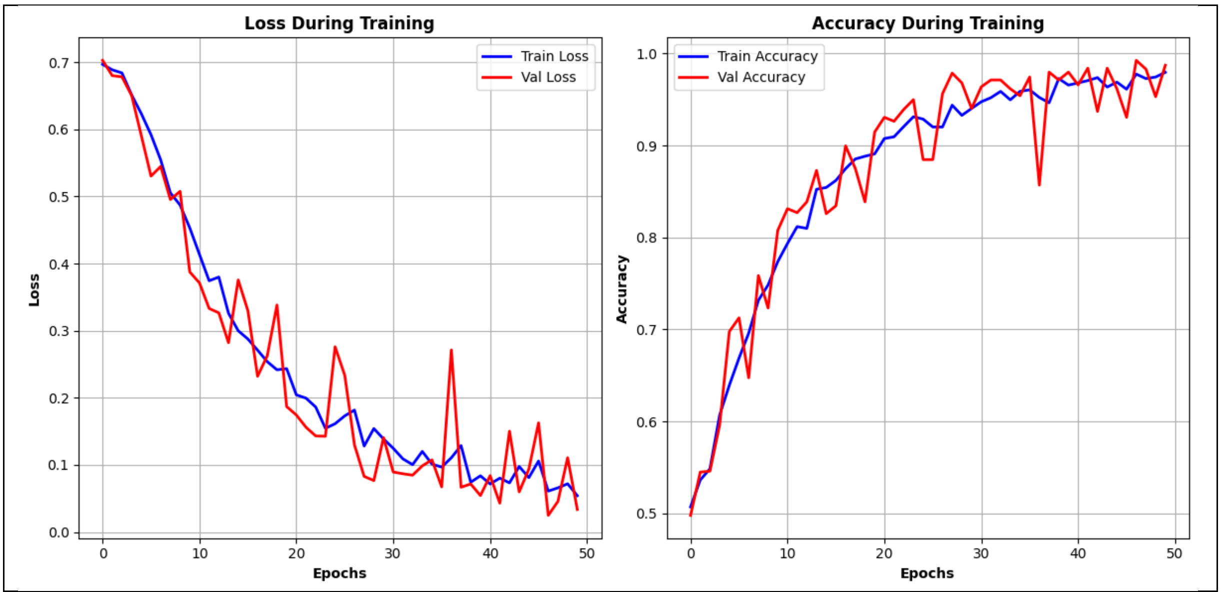
Figure 5: Confusion Matrix for machine learning models

4.2 Deep Learning Comparative Analysis

Deep Learning comparative analysis compares the performance of all neural networks described in the methodology section. The section describes the comparison of all algorithms across various comparative metrics. It compares all the different deep learning algorithms namely Artificial neural networks, Dense neural networks, and Convolutional neural networks.



(a) Artificial Neural Network



(b) Dense Neural Network

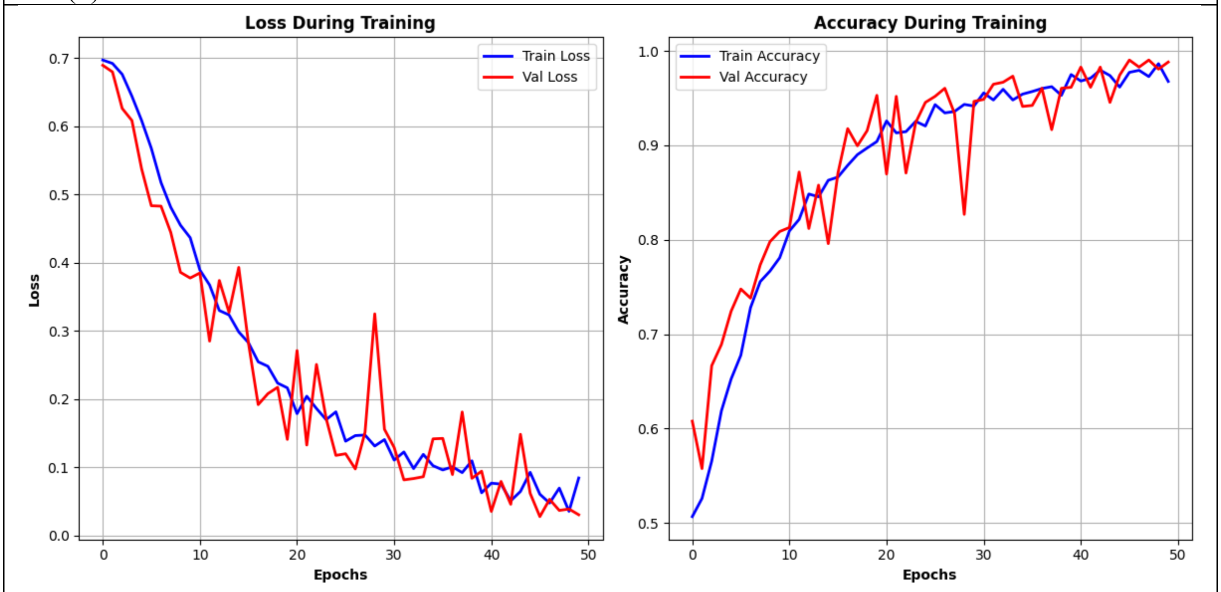


Figure 6: Training and Validation Curves

Figure 6: Training and Validation Curves Describes the comparison of all algorithms across various comparative graphical representations. Allowing a view of global training and its validation performance. It compares all the different deep learning algorithms namely Artificial neural networks, Dense neural networks, and Convolutional neural networks.

Table 5: Deep Learning Analysis

MODEL	ACCURACY	COHEN'S KAPPA	JACCARD SIMILARITY	BENIGN PRECISION	BENIGN RECALL	BENIGN F1-SCORE	MALIGNANT PRECISION	MALIGNANT_RECALL	MALIGNANT F1-SCORE
ANN	0.996154	0.992308	0.992337	0.99914	0.993162	0.996142	0.993203	0.999145	0.996165
DNN	0.985844	0.971687	0.972078	0.998902	0.972727	0.985641	0.973479	0.998933	0.986042
CNN	0.991987	0.983974	0.984102	0.989362	0.994652	0.992	0.994635	0.989328	0.991974

As shown in Table 5: Deep Learning Analysis Model ANN achieved an accuracy of 0.996154 with a Cohen's Kappa of 0.992308 and Jaccard Similarity of 0.992337.

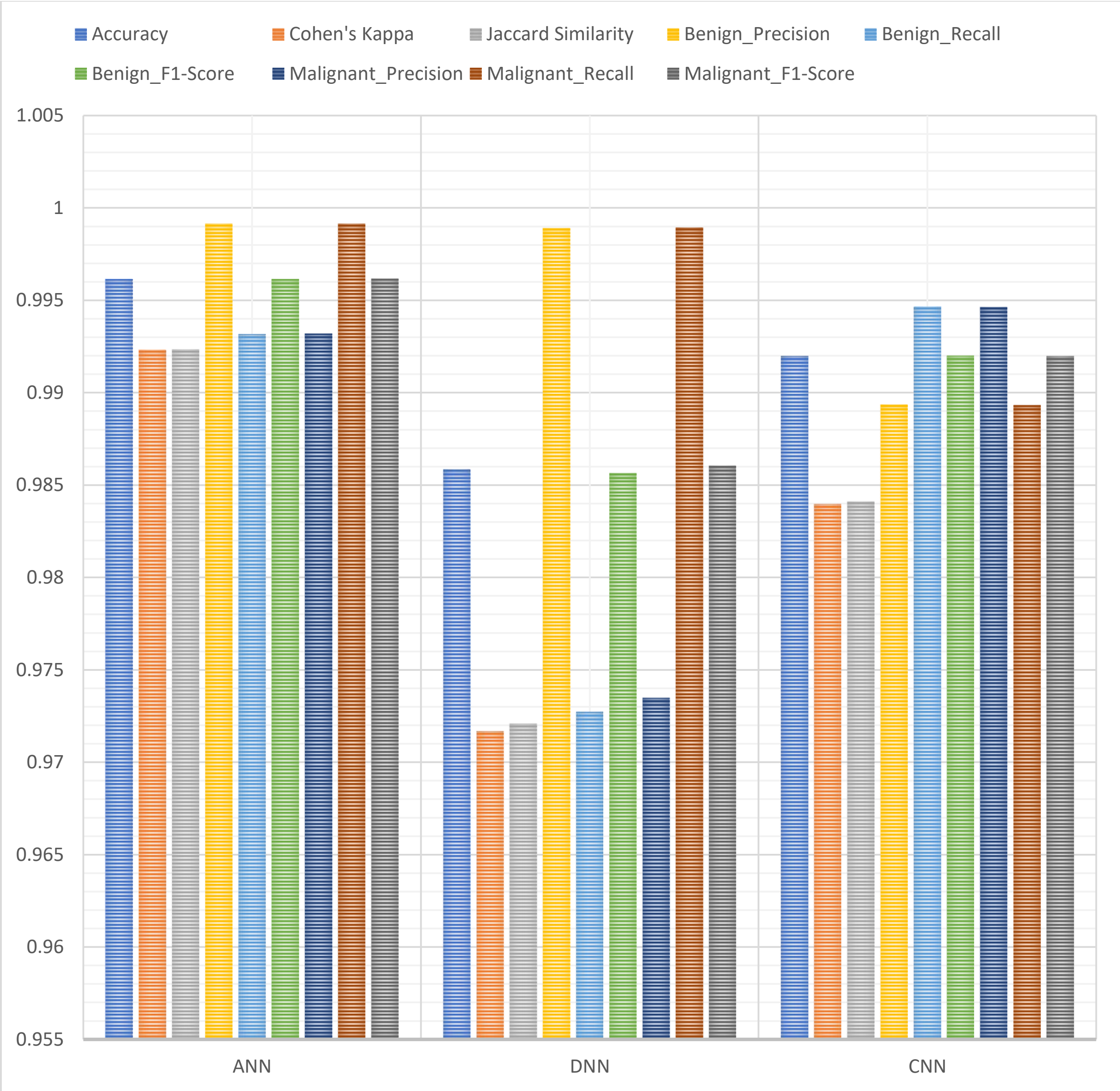
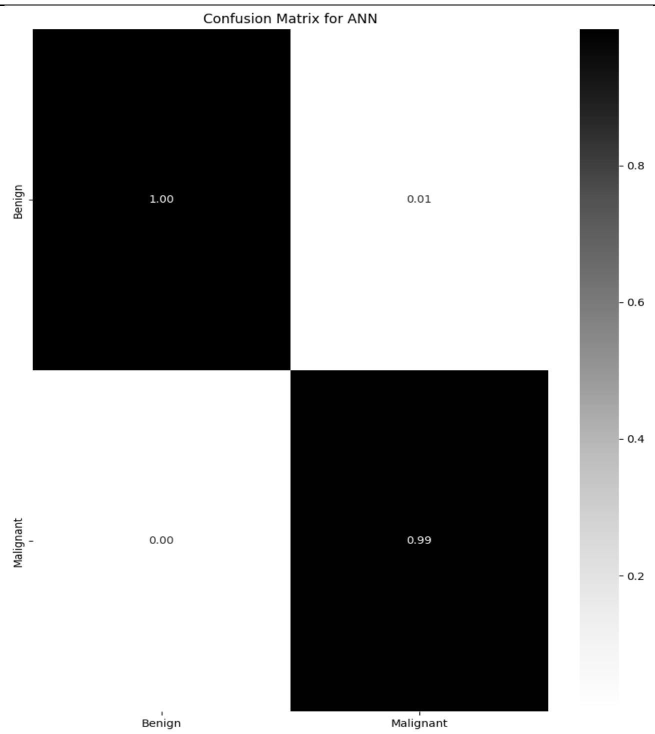
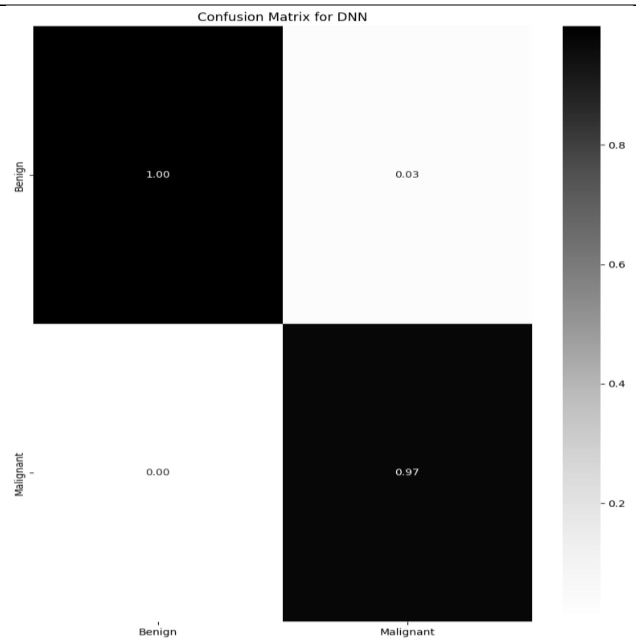


Figure 7: Deep learning Model Visualisation

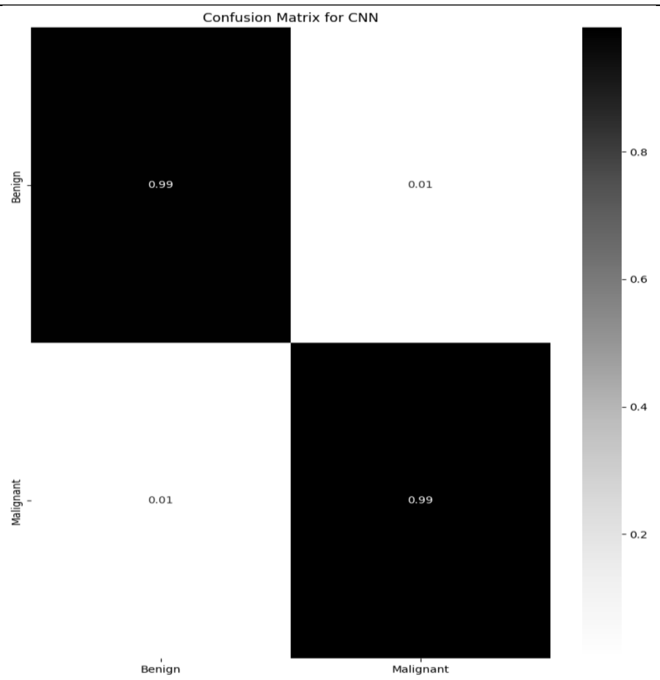
Figure 7: Deep learning Model Visualisation Describes the comparison of all algorithms across various comparative metrics. Allowing a view of global metrics and their performance. It compares all the different deep learning algorithms namely Artificial neural networks, Dense neural networks, and Convolutional neural networks.



(a) Artificial Neural Network



(b) Dense Neural Network



(c) Convolutional Neural Network

Figure 8: Confusion Matrix for deep learning models

Figure 8: Confusion Matrix for deep learning models the showcases the Confusion matrix for all the different deep learning algorithms namely Artificial neural network, Dense neural network, and Convolutional neural network

4.3 RCANN or Residual Convolutional Attention Neural Network

RCANN showcases the performance of RCA neural networks described in the methodology section.

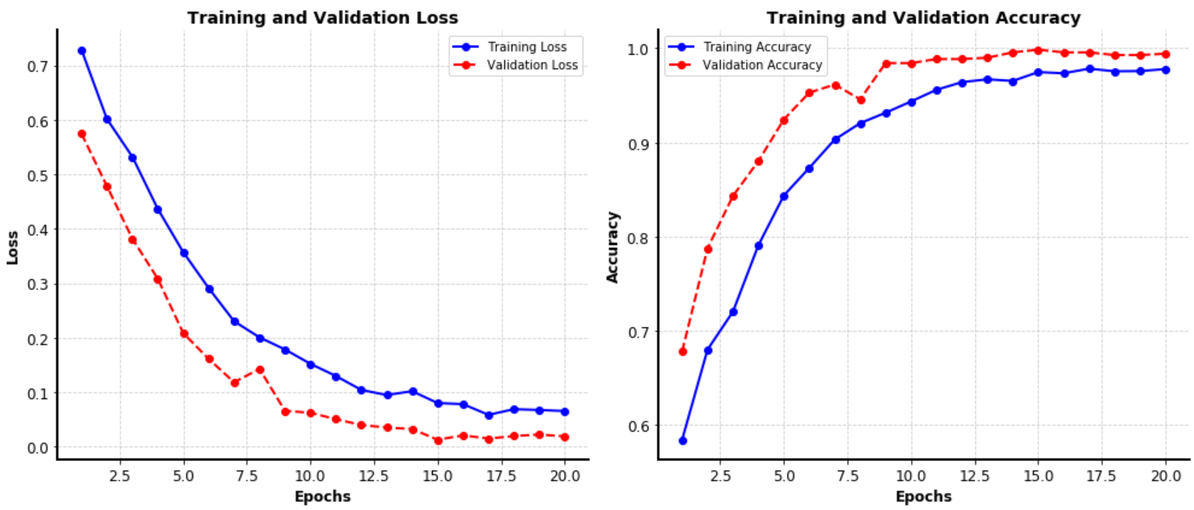


Figure 9: RCANN Training and Validation Curve

Figure 9: RCANN Training and Validation Curve Showcases the graphical representations. Allowing a view of global training and its validation performance.

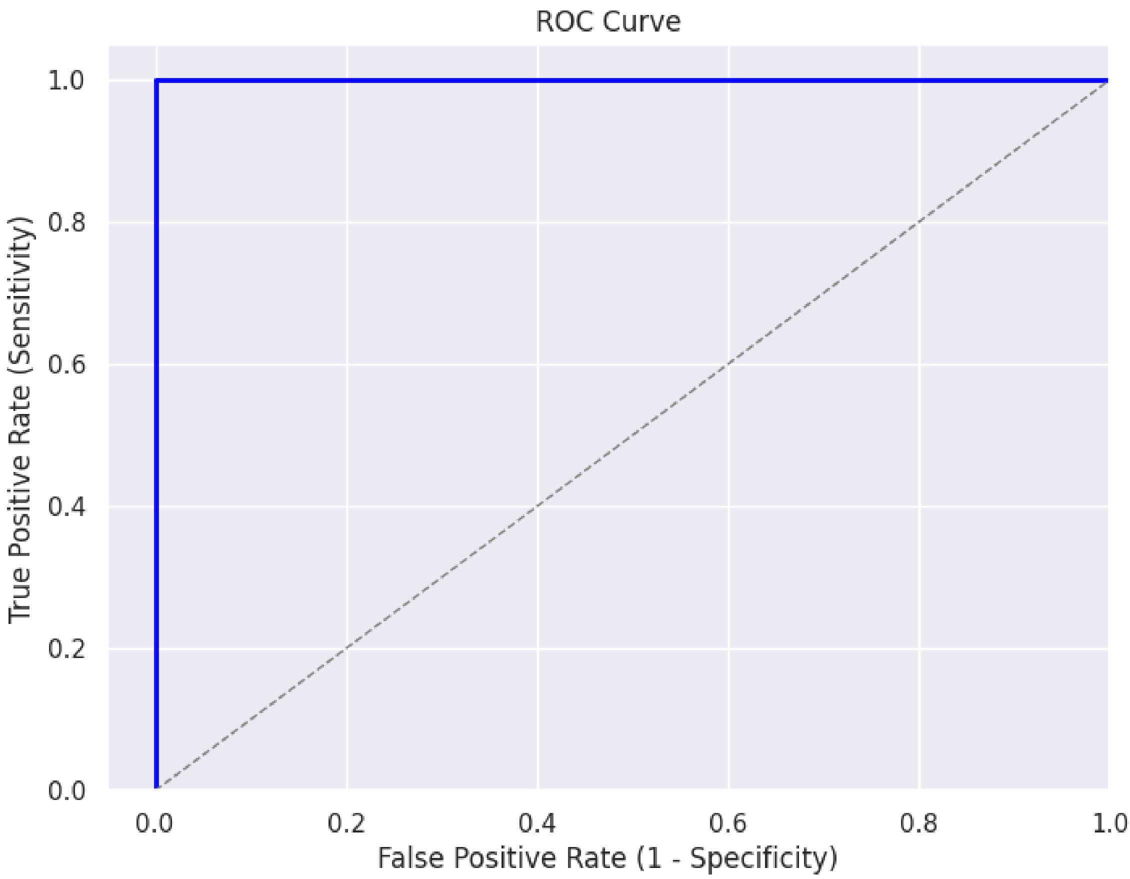


Figure 10: RCANN ROC Curve

Figure 10: RCANN ROC Curve Showcases the ROC curve and area under the curve for the RCANN.

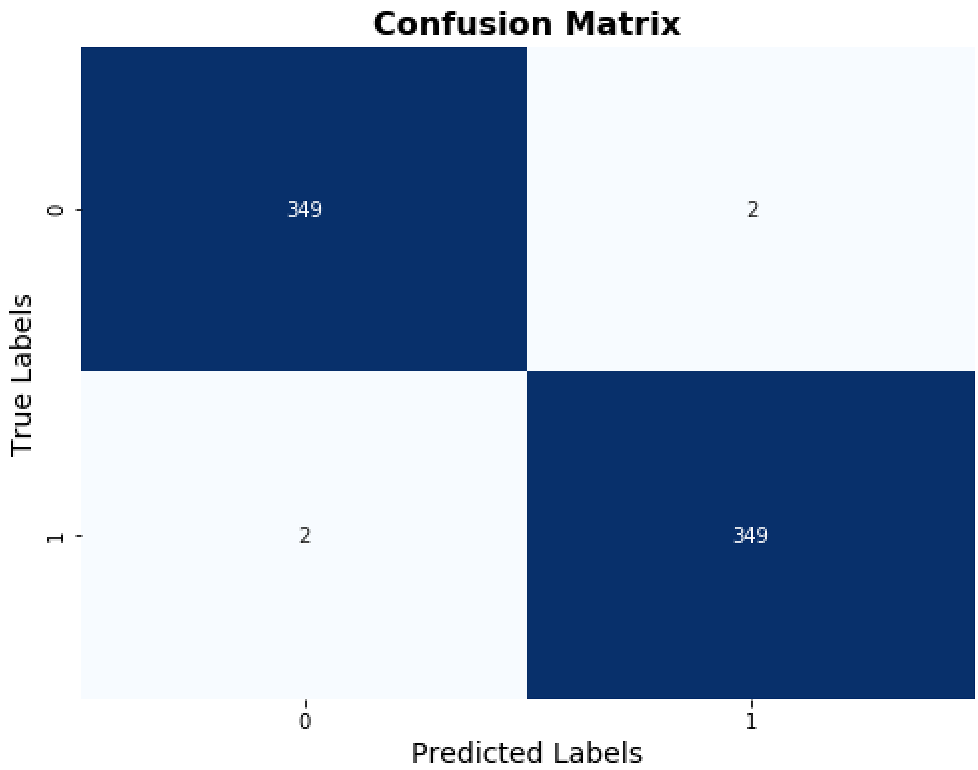


Figure 11: RCANN Confusion Matrix

Figure 11: RCANN Confusion Matrix Showcases the Confusion Matrix for the RCANN.

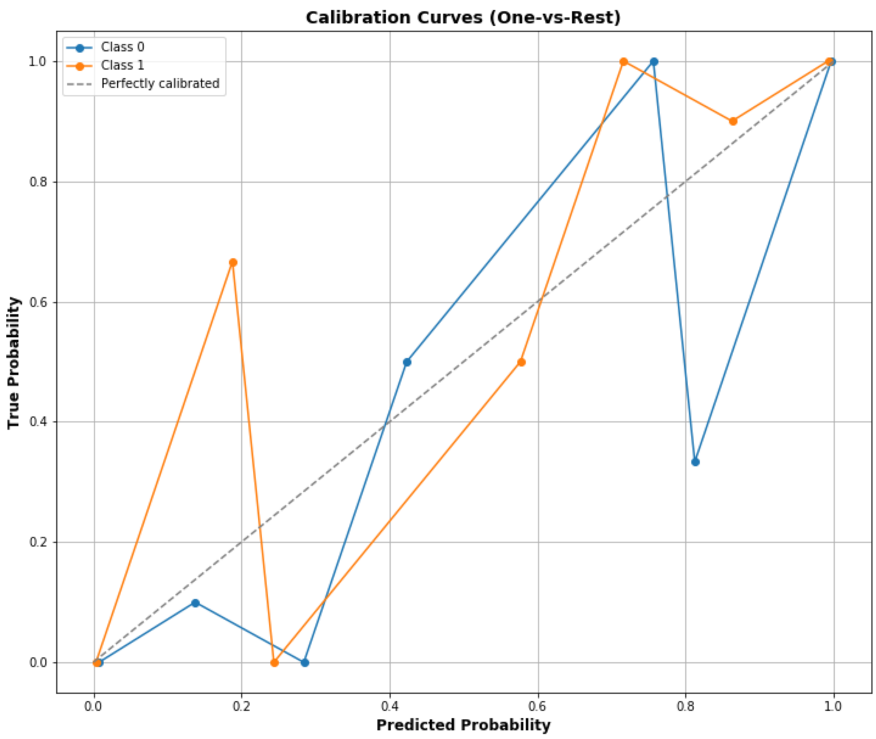


Figure 12: RCANN precision-recall curve

Figure 12 RCANN precision-recall curve, shows a Precision-Recall curve for a binary classification task, with Class 0 (blue) and Class 1 (orange) plotted along with a reference line for perfect calibration

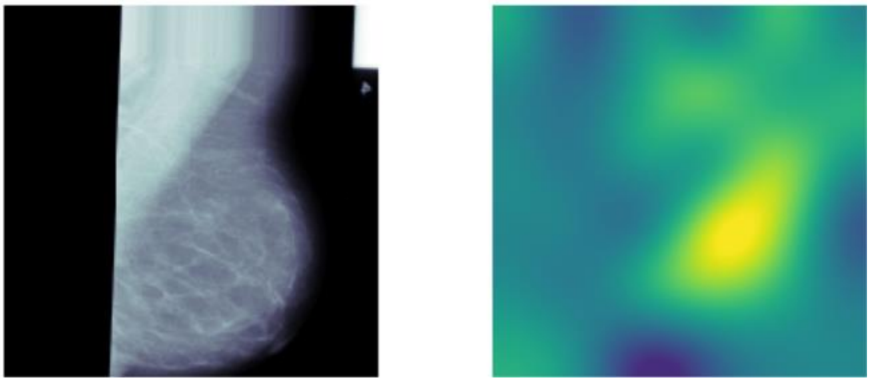
(Gray dashed). It illustrates the trade-off between precision and recall at various thresholds for both classes.

Table 6: RCANN metric of performance

METRIC	RCANN
ACCURACY	0.99987
COHEN'S KAPPA	0.99997
MATTHEWS CORRELATION COEFFICIENT	0.99946
LOG LOSS	0.02
BENIGN PRECISION	0.99158
BENIGN RECALL	0.99220
BENIGN F1-SCORE	0.99009
MALIGNANT PRECISION	0.99343
MALIGNANT RECALL	0.99341
MALIGNANT F1-SCORE	0.99008

Table 6: RCANN metric of performance Showcase RCANN achieved an accuracy of 0.99987, Cohen's Kappa of 0.99997, Matthews Correlation Coefficient of 0.99946, and a Log Loss of 0.02. For benign cases, it achieved a precision of 0.99158, recall of 0.99220, and F1-Score of 0.99009. For malignant cases, it achieved a precision of 0.99343, recall of 0.99341, and F1-Score of 0.99008.

4.4 Explainable AI



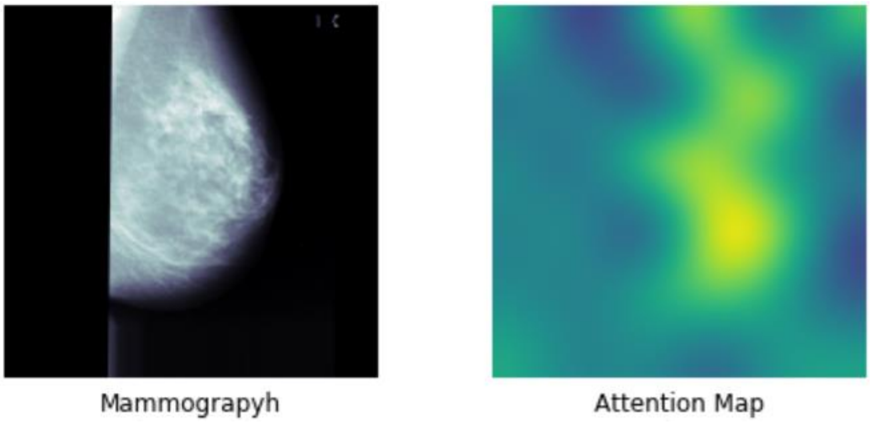


Figure 13: Malignant Samples Attention Network Output

Figure 13: Malignant Samples Attention Network Output Displays the malign tumors with its attention network output, which highlights the key regions of concern in the breast tissue.

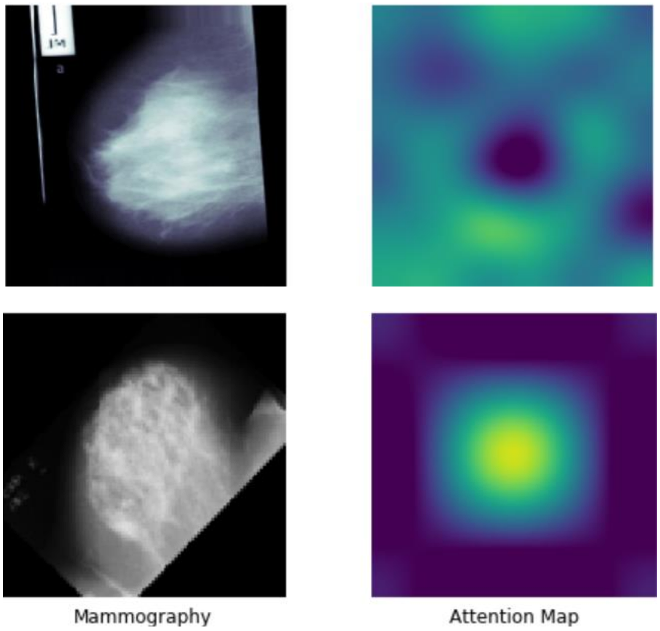


Figure 14: Benign Sample Attention Network Output

Figure 14: Benign Sample Attention Network Output Shows the benign tumor and its corresponding attention network output, where the attention map indicates no significant abnormalities compared to the malignant case. This allows an extension to the explainable artificial intelligence domain.

The attention maps help identify the region of interest within the selected tissues of the monogamic breast cells. The explainable AI the primary aim is to make deep neural sciences more visible and interpretable. This allows medical image analysis to be more user-interactive and prevents false positives from being redirected into the system as predictions. The images Figure 13: Malignant Samples Attention Network Output and Figure 14: Benign Sample Attention Network Output Showcases the exemplary examples of attention methods to achieve explainable AI. These maps illustrate the exact parts of the mammography that the model deems significant for its predictions, offering insight into the program's decision-making process.

4.5 State-of-the-Art Comparison

The section compares the state-of-the-art comparison with the RCANN model with the Table 7: SOTA comparison. The Model RCANN outperforms the aforementioned models, RCANN achieved an accuracy of 0.99987, Matthews Correlation Coefficient of 0.99946, and a Log Loss of 0.02.

Table 7: SOTA comparison

Methods	*Accuracy*	*Precision*	*F1 Score*
EKNN	96.1	96.0	96.4
FKNN	94.8	94.7	95.1
WKNN	94.9	94.8	95.2
LSVM	94.1	94.1	94.5
QSVM	94.8	94.8	95.2
CSVM	94.9	95.0	95.3
MKSVM	96.1	95.9	96.3
MNN	94.9	94.9	95.2
TNN	94.9	94.9	95.3
BLNN	95.5	95.6	95.8
ELM	87.1	85.7	N/A
KELM	92.4	91.2	N/A
MRNN	93.3	94.5	N/A
FCNN	95.0	95.0	N/A
RCANN (proposed)	**99.987**	**99.343**	**99.008**

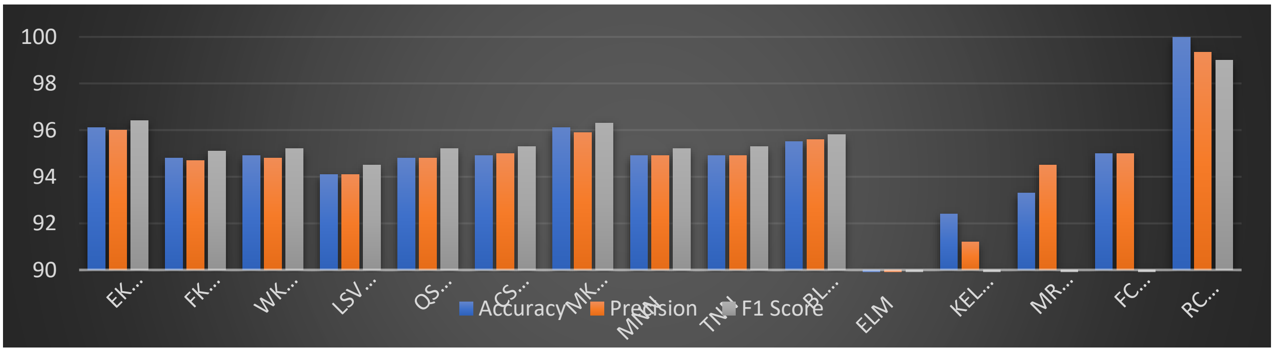


Figure 15: State-of-the-art comparison bar plot

The Figure 15: State-of-the-art comparison bar plot Compares the accuracy, precision, and F1 score for EKNN, FKNN, WKNN, LSVM, QSVM, CSVM, MKSVM, MNN, TNN, BLNN, ELM, KELM, MRNN, FCNN, and RCANN (proposed) techniques, with RCANN surpassing all.

5 Conclusion

The study proposes a detailed understanding of breast cancer, the current limitations for automated detection of breast cancer, and research gaps surrounding the detection. The research builds upon the research gap across the literature review and proposes the RCANN module a residual convolutional attention neural network. The network is thoroughly compared across machine learning techniques like random forest, decision tree, voting classifier, and adaptive boosting. Further to have a detailed simulation of deep neural architecture it is compared with artificial neural networks, dense neural networks, and convolutional neural networks.

The proposed RCANN has been compared across various performance metrics such as precision, recall, F1 score, accuracy, sensitivity, specificity, and support. The metrics are further visualized by bar plot, confusion matrix, and ROC-AUC curves. The proposed RCANN model is also capable of deducing the attention maps to create strong kernel views aiding towards explainable AI, this is empirical for the research towards the detection and use of AI as an aid towards medical image processing in breast cancer sampling images.

5.1 Future Scope

The research on breast cancer and medical image processing is an ever-evolving segment however based on this research future researches need to contribute towards the segmentation and localization of the medical image blob allowing for the staggering transfer of the information from passive breast cancer detection to active breast cancer simulation localization mapping. This contour framework will further allow the enhancement of AI dependency and mobility of cancer in the early stages. The schematic prediction can be further nuanced by federated and mobile-based deep learning for smaller computational algorithms.

6 REFERENCES

[1]. Marks, J.S.; Lee, N.C.; Lawson, H.W.; Henson, R.; Bobo, J.K.; Kaeser, M.K. Implementing recommendations for the early detection of breast and cervical cancer among low-income women. Morb. Mortal. Wkly. Rep. Recomm. Rep. 2000, 49, 35–55.

[2]. Schueller, G.; Schueller-Weidekamm, C.; Helbich, T.H. Accuracy of ultrasound-guided, large-core needle breast biopsy. Eur. Radiol. 2008, 18, 1761–1773. https://sci-hub.se/10.1007/s00330-008-0955-4

[3]. Ferlay, J.; Colombet, M.; Soerjomataram, I.; Parkin, D.M.; Piñeros, M.; Znaor, A.; Bray, F. Cancer statistics for the year 2020: An overview. Int. J. Cancer 2021, 149, 778–789. https://sci-hub.se/10.1002/ijc.33588

[4]. Lei, S.; Zheng, R.; Zhang, S.; Wang, S.; Chen, R.; Sun, K.; Zeng, H.; Zhou, J.; Wei, W. Global patterns of breast cancer incidence and mortality: A population-based cancer registry data analysis from 2000 to 2020. Cancer Commun. 2021, 41, 1183–1194.

[5]. Breast Cancer Detection in Mammography Images: A CNN-Based Approach with Feature Selection written by Zahra Jafari and Ebrahim Karami

[6]. Sun, L.; Wang, J.; Hu, Z.; Xu, Y.; Cui, Z. Multi-view convolutional neural networks for mammographic image classification. IEEE Access 2019, 7, 126273–126282.

[7]. Amrane, M., Oukid, S., Gagaoua, I. & Ensari, T. Breast cancer classification using machine learning. In 2018 Electric Electronics, Computer Science, Biomedical Engineering's Meeting (EBBT) 1–4 (IEEE, 2018). https://sci-hub.se/10.1109/ebbt.2018.8391453

[8]. Wang, Z. et al. Breast cancer detection using extreme learning machine based on feature fusion with CNN deep features. IEEE Access 7, 105146–105158 (2019).

[9]. Khan, H. N., Shahid, A. R., Raza, B., Dar, A. H. & Alquhayz, H. Multi-view feature fusion based four views model for mammogram classification using convolutional neural network. IEEE Access 7, 165724–165733 (2019).
[10]. Saber, A., Sakr, M., Abo-Seida, O. M., Keshk, A. & Chen, H. A novel deep-learning model for automatic detection and classification of breast cancer using the transfer-learning technique. IEEE Access 9, 71194–71209 (2021).
[11]. Omonigho, E. L., David, M., Adejo, A. & Aliyu, S. Breast cancer: Tumor detection in mammogram images using modified AlexNet deep convolution neural network. In 2020 International Conference in Mathematics, Computer Engineering and Computer Science (ICMCECS) 1–6 (IEEE, 2020). https://sci-hub.se/10.1109/icmcecs47690.2020.240870
[12]. Singh, R. et al. Imbalanced breast cancer classification using transfer learning. IEEE/ACM Trans. Comput. Biol. Bioinform. 18(1), 83–93 (2020). https://sci-hub.se/10.1109/tcbb.2020.2980831
[13]. Liu, K., Kang, G., Zhang, N. & Hou, B. Breast cancer classification based on fully-connected layer first convolutional neural networks. IEEE Access 6, 23722–23732 (2018).
[14]. Sarosa, S. J. A., Utaminingrum, F. & Bachtiar, F. A. Mammogram breast cancer classification using gray-level co-occurrence matrix and support vector machine. In 2018 International Conference on Sustainable Information Engineering and Technology (SIET) 54–59 (IEEE, 2018). https://sci-hub.se/10.1109/siet.2018.8693146
[15]. Zebari, D. A., Zeebaree, D. Q., Abdulazeez, A. M., Haron, H. & Hamed, H. N. A. Improved threshold based and trainable fully automated segmentation for breast cancer boundary and pectoral muscle in mammogram images. IEEE Access 8, 203097–203116 (2020).
[16]. Liu, M. et al. A deep learning method for breast cancer classification in the pathology images. IEEE J. Biomed. Health Inform. 26(10), 5025–5032 (2022). https://www.researchgate.net/publication/361691358_A_Deep_Learning_Method_for_Breast_Cancer_Classification_in_the_Pathology_Images
[17]. Hamed, G., Marey, M., Amin, S. E. & Tolba, M. F. Automated breast cancer detection, and classification in full field digital mammograms using two full and cropped detection paths approach. IEEE Access 9, 116898–116913 (2021).
[18]. Zhang, X. et al. Classification of whole mammogram and tomosynthesis images using deep convolutional neural networks. IEEE Trans. Nanobiosci. 17(3), 237–242 (2018). https://sci-hub.se/10.1109/tnb.2018.2845103
[19]. Rastghalam, R. & Pourghassem, H. Breast cancer detection using MRF-based probable texture feature and decision-level fusion-based classification using HMM on thermography images. Pattern Recognit. 51, 176–186 (2016). https://sci-hub.se/10.1016/j.patcog.2015.09.009
[20]. Liu, J., Feng, J. & Gao, X. Fault diagnosis of rod pumping wells based on support vector machine optimized by improved chicken swarm optimization. IEEE Access 7, 171598–171608 (2019).
[21]. Wang, J. et al. Improvement and application of chicken swarm optimization for constrained optimization. IEEE Access 7, 58053–58072 (2019).
[22]. Deore, B. & Bhosale, S. Hybrid optimization enabled robust CNN-LSTM technique for network intrusion detection. IEEE Access 10, 65611–65622 (2022).
[23]. Hayum, A. A., Jaya, J., Sivakumar, R., & Paulchamy, B. (2024). An efficient breast cancer classification model using bilateral filtering and fuzzy convolutional neural network. Scientific Reports, 14(1), 6290.
[24]. Chakravarthy, S., Nagarajan, B., Kumar, V. V., Mahesh, T. R., Sivakami, R., & Annand, J. R. (2024). Breast Tumor Classification with Enhanced Transfer Learning Features and Selection Using Chaotic Map-Based Optimization. International Journal of Computational Intelligence Systems, 17(1), 1-18.