Enhancing Breast Cancer Classification Accuracy through Transfer Learning with DenseNet121: A Comparative Study with Conventional CNN Models

Blessing Olorunfemi *olorunfemib@run.edu.ng*
Department of Computer Science,
Redeemer's University,
Ede, Nigeria

Adewale Ogunde *ogundea@run.edu.ng*
Department of Computer Science,
Redeemer's University,
Ede, Nigeria

Adenike Adeniji-Sofoluwe *nikesofoluwe@gmail.com*
Department of Radiology,
College of Medicine,
University of Ibadan, Nigeria

Bosede Oguntunde *oguntunden@run.edu.ng*
Department of Computer Science,
Redeemer's University,
Ede, Nigeria

Samson Arekete *areketes@run.edu.ng*
Department of Computer Science,
Redeemer's University,
Ede, Nigeria

Alex Pearson *alexander.Pearson@uchicagomedicine.org*
Department of Medicine,
University of Chicago,
Chicago, USA

Funmilayo Olopade *folopade@bsd.uchicago.edu*
Center for Clinical Cancer
Genetics and Global Health,
University of Chicago;
Chicago, USA
&Department of Medicine,
University of Chicago,
Chicago, USA

Benjamin Aribisala *aribisala@uchicago.edu*
Department of Computer Science,
Lagos State University, Nigeria
&Center for Clinical Cancer
Genetics and Global Health,
University of Chicago;
Chicago, USA

Abstract

Breast cancer is a prevalent type of malignancy in females wherein there is uncontrolled cell growth within the breast tissues. Proper identification and classification are the basis for effective

treatment and management. There has been potential in increasing classification accuracy as well as support for early diagnosis through more recent advancements with deep learning models, particularly when utilized in medical imaging. This research aims to enhance the precision of breast cancer classification by comparing deep learning model performance. Python and deep learning frameworks were employed in developing and comparing models for breast cancer classification using the Curated Breast Imaging Subset of the Digital Database for Screening Mammography (CBIS-DDSM) dataset, which includes Digital Imaging and Communications in Medicine (DICOM) mammography images obtained through Kaggle. The data was quality-assured and made uniform. A conventional Convolutional Neural Network (CNN) was first applied for binary classification. Transfer learning was implemented with the DenseNet121 model that was pre-trained on ImageNet to improve performance. Layers of the model were frozen, and classification layers were included as custom. Fine-tuning was accomplished by unfreezing certain layers to enhance the ability of the model to discriminate between malignant and benign cases. The conventional CNN model achieved accuracy of 51.87%, weighted F1-score of 0.35, precision of 0.27, and recall of 0.52. Following transfer learning with DenseNet121, accuracy was improved to 71%, weighted F1-score of 0.71, Specificity of 0.83, Sensitivity of 0.61 and AUC of 0.7. Fine-tuning resulted in an end accuracy of 88%, with weighted F1-score, Sensitivity of 0.87, Specificity of 0.82, precision at 0.87 and Area Under the Curve (AUC) at 0.85.This study highlights the effectiveness of DenseNet121 combined with transfer learning for improving breast cancer classification accuracy using DICOM images from the CBIS-DDSM dataset, contributing to more reliable early detection and treatment strategies.

Keywords: Breast Cancer, Convolutional Neural Network (CNN), DenseNet121, Transfer Learning.

1. INTRODUCTION

Breast cancer is the most prevalent type of malignancy found globally, notably in women (Yiallourou, 2023; Smolarz et al., 2022; Cuthrell & Tzenios, 2023). It is defined by the uncontrollable reproduction of cellular tissues found inside the breast tissues. The condition can develop into tumors (Wang et al., 2024; Akinpelu et al., 2024; Akl & Ahmed, 2024; Ojo et al., 2025; Hong et al., 2025). The earlier the detection of the malignant cells present in the body, the more promising the treatment is for the patient (Xiong et al., 2025). Nonetheless, the condition may be experienced by the patient without symptoms at all until the point of malignancy.

Despite its importance as a diagnostic imaging tool, mammography image analysis is a time-consuming process that requires human expertise, resulting in possible inaccuracies (Dave et al., 2025; Santos et al., 2024). In an effort to overcome these inconveniences, Computer-Aided Diagnostic (CAD) software was conceived to support and help radiologists identify and classify cancer accurately (Hussain et al., 2024). This aid utilizes machine learning, specifically deep learning, to evaluate images.

Convolutional Neural Networks (CNNs) are one of the most popular deep learning models for the assessment of medical images because of their capability to detect complicated patterns in images (Zangana et al., 2024; Mienye et al., 2025; Takahashi et al., 2024; Manjunatha & Mahendra, 2024). However, regular CNN models have limitations with complicated images like the DICOM mammography images (Sharafaddini et al., 2024; Salehi et al., 2023). However, the need to train models using a large number of images and prevent vanishing gradients has promoted the concept of transfer learning, which can work efficiently even while handling fewer labeled images (Chutia et al., 2024). DenseNet121, one of the CNN models designed using the concept of dense connectivity, turns out to be one of the most accurate models for such applications and avoids complications such as vanishing gradients (Blahová et al., 2025).

The current study examines the capability of the DenseNet121 architecture with the use of transfer learning for the classification task of breast cancer based on mammography images. The

results derived from the proposed method are then compared to a basic CNN architecture to ascertain enhancements made within the context of accuracy, recall, precision, and overall classification performance.

2. LITERATURE REVIEW

The development of Deep Learning algorithms that perform significantly better than traditional ones has completely altered the landscape of research work done in the classification of mammography and histopathology images. Traditional approaches involved the use of classical Machine Learning models such as Support Vector Machines (SVMs). The model used manually designed features, including texture and mass properties. For instance, the work done by Falconi et al. in 2020 highlighted the applicability of the SVM approach to the DDSM database, where they extracted a total of 181 features, resulting in an AUC of 0.805. The approach showed promise but was limited by the need to design features manually, which was not particularly successful in the case of medical images that have complex properties. The development in Deep Learning, specifically Convolutional Neural Networks, introduced a new approach that exploited automatic feature extraction. The adoption of transfer learning approaches was also early and greatly appreciated in the field of medical imaging, as annotated datasets are limited in these fields. For example, learning was transferred using ResNet50 and MobileNet by Ansar et al. (2020); accuracies of 78.4% and 74.3% were consequently obtained for the classification task on the MIAS dataset. The same was adopted by Prusty et al. (2022) with the use of MobileNet, and accuracies of 86.8% and 74.5% were obtained for the DDSM and CBIS-DDSM datasets, respectively. Additionally, accuracy of 81.99% was obtained by Sharma et al. (2022) using VGG16 on the MIAS dataset, showing its applicability despite its limitations.

However, aside from mammography, the effectiveness of transfer learning has also been validated in other applications within the field of medical imaging. For example, Kaur and Mahajan (2025) were able to reach a recognition accuracy of 98.53% for the classification of brain tumors through the use of ResNet152 combined with GoogleNet, exemplifying the applicability and efficacy of CNNs across various tasks and datasets. Another example is given by Zahra et al. (2023), who was able to reach an accuracy rate of 98.4% utilizing the features offered by the use of DenseNet, noting the significant impact and need for image preprocessing to improve performance within a specific class. Elkorany et al. (2023) further improve these findings by combining the features gathered from multiple CNNs and then utilizing other machine learning algorithms such as KNN, SVM, as well as RF to gather an accuracy rate of 94.5% within the MIAS image repository. This was further supported by Chakravarthy et al. (2024).

More recent works aim to optimize CNN architectures and further improve models. Meenalochini and Ramkumar (2024) used EfficientNet-B4 and reached 98.46% for INbreast. Laishram and Rabidas (2024) used image preprocessing methods and implemented Genetic Algorithms, and as a result, they enhanced classification accuracy to 83.5% on INbreast. Lee et al. (2017) implemented a combination of machine learning and deep learning approaches and reached 83.61% on mini-MIAS. Transfer learning models utilizing DenseNet121 attracted recent attention. Laishram and Rabidas (2024) implemented DenseNet121 on BreakHis histopathological images and reached 96.09% on 100x magnified images. Bello et al. (2024) further optimized DenseNet121 for skin lesion classification. They implemented additional dense layers with leaky ReLU activation and outperformed EfficientNetB0, ResNet34, and VGG16. Ahad et al. (2024) illustrated that although CNNs perform well independently, their combination can further improve prediction and reached 99.94% using an ensemble of six models including DenseNet121, ResNet18, and MobileNetV2. Improved preprocessing and data augmentation methods were proven to further improve models performance. Seneng et al. (2025) utilized a combination of techniques such as CLAHE, Median Filtering, and Wavelet Transform and found VGG19 to reach 98.04% accuracy with DWT-enhanced mammography images, although issues of imbalance were shown to exist. Chugh et al. (2024) made a comparison between transfer learning and traditional approaches to machine learning. Their study found that transferred CNNs (MobileNet, ResNet50, and DenseNet169) performed well and had around 97% accuracy, learning around 4% better than SVM and RandomForest. Lastly, Taifi et al. (2025) adjusted DenseNet121,

DenseNet201, and MobileNetV2 models by freezing some of the initial convolutional blocks and used GELUs as activation functions. Their optimized models showed 97% to 99.6% accuracy on the MIAS, INbreast, and DDSM databases, and that improvements were statistically proven using paired t-tests.

Despite the significant achievements of using transfer learning through advanced CNN architectures like DenseNet, VGG, ResNet, and EfficientNet for the diagnosis of breast cancer, some research loopholes still need to be addressed. Current studies mainly focused on the comparison of various pre-trained architectures of deep learning without giving much importance to the comparison of the aforementioned architectures with conventional CNN architectures. Moreover, although various datasets including BreakHis and MIAS have achieved significant attention from the research community, little attention has been paid to the mammography CBIS-DDSM dataset, although its diagnostic challenges are higher. In addition to these limitations, although some studies through architectural improvement and ensemble methods demonstrated near-perfect accuracy in their respective results, the proposed models featured complex architectures that reduce their practical applications in real-world scenarios. In the proposed approach, an exhaustive comparison between the baseline CNN architecture and the transfer learning method using the DenseNet121 model for the mammography CBIS-DDSM dataset would address these limitations. It would clearly reveal the significant improvements achieved through transfer learning to prove the authenticity and importance of DenseNet121 in breast cancer detection. Table 1 shows the overall summary of the prior art works in the proposed area.

S/N	*Author(s), Year*	*Identified Gap*	*Method*	*Result*	*Limitation*
1	Falconi et al. (2020)	Classical ML models required manual feature design	SVM with 181 handcrafted features on DDSM	AUC = 0.805	Dependent on manual feature extraction; limited ability to capture complex patterns
2	Ansar et al. (2020)	Limited exploration of deep learning in mammogram classification	Transfer learning using ResNet50 and MobileNet on MIAS	78.4% and 74.3% accuracy	Small dataset size; moderate performance
3	Prusty et al. (2022)	Need for improved accuracy across datasets	Pre-trained MobileNet on DDSM and CBIS-DDSM	86.8% and 74.5% accuracy	Dataset variability affected performance
4	Sharma et al. (2022)	Traditional CNNs limited by feature depth	VGG16 on MIAS dataset	81.99% accuracy	Limited scalability across datasets
5	Kaur and Mahajan (2025)	Applications beyond mammograms underexplored	ResNet152 and GoogLeNet on brain tumor MRI	98.53% recognition rate	Performance may not generalize to heterogeneous datasets
6	Zahra et al. (2023)	Impact of preprocessing	DenseNet with preprocessing	98.4% accuracy	Dataset-specific; generalizability

		on DL models unclear			unclear
7	Elkorany et al. (2023)	Single CNN models prone to limitations	Multi-CNN feature extraction + KNN, SVM, RF	94.5% accuracy	Increased complexity; interpretability challenges
8	Chakravarthy et al. (2024)	Hybrid feature pipelines underexplored	Inception-V3, ResNet50, AlexNet features + SVM	98% accuracy	High computational overhead
9	Meenalochini and Ramkumar (2024)	Need to balance CNN depth and efficiency	EfficientNet-B4 on INbreast	98.46% accuracy	Requires high computational resources
10	Laishram and Rabidas (2024)	Feature optimization limited	Genetic algorithms with preprocessing	83.5% accuracy	Lower performance than DL baselines
11	Lee et al. (2017)	Pure DL models struggled with dataset variability	Hybrid ML + DL on mini-MIAS	83.61% accuracy	Modest performance; dataset constraints
12	Potsangbam and Devi (2024)	DenseNet121 underexplored in histopathology	DenseNet121 on BreakHis dataset	96.09% accuracy (100× magnification)	Limited to histopathological images
13	Bello et al. (2024)	Need to optimize DenseNet121	DenseNet121 with added dense layers and LeakyReLU	Outperformed EfficientNetB0, ResNet34, VGG16	Increased model complexity
14	Ahad et al. (2024)	Individual CNNs underperformed	Ensemble of six CNNs including DenseNet121	99.94% accuracy	Computationally expensive; reduced interpretability
15	Seneng et al. (2025)	Role of preprocessing in CNNs unclear	VGG19 with CLAHE, median filtering, wavelets	98.04% accuracy	Dataset imbalance issues
16	Chugh et al. (2024)	Limited comparison between classical ML and CNNs	MobileNet, ResNet50, DenseNet169 vs. SVM/RF	~97% CNN accuracy; ~4% higher than ML	Requires large datasets

17	Taifi et al. (2025)	Need for improved fine-tuning strategies	Modified DenseNet121, DenseNet201, MobileNetV2 with GELU	97–99.6% accuracy on MIAS, INbreast, DDSM	High training and tuning complexity

TABLE 1: Summary of Related Works on Breast Cancer Classification

3. MATERIALS AND METHODS

In order to determine the effectiveness of deep learning models for the classification of breast cancer images, the study adopts an experimental research design. The study proposes a comparative study of a regular Convolutional Neural Network (CNN) model versus a transfer learning model utilizing a DenseNet121 model. In relation to the above, a CNN model and a transfer learning model utilizing a DenseNet121 model were used for the classification of images from the CBIS-DDSM database to yield better results for the diagnosis of breast cancer. A step-by-step process of the study is presented in Figure 1.

Starting with the CBIS-DDSM dataset, which is downloaded from the Kaggle platform and initially converted using various preprocessing methods like rotation, augmentation, normalization, and scaling to make it ready for the process of training models. In an attempt to robustly train and test the models, the dataset is split equally into the training and testing data sets. In an effort to construct a standard for comparison, the initial model built is a standard CNN binary classifier (benign vs. malignant) model. Subsequently, the DenseNet121 model, an advanced deep learning model well-known for being remarkably proficient at the task of extracting relevant features from the given images, is introduced. On the back of the pre-trained image net weights, the DenseNet121 model begins with the freezing of the initial layers and the subsequent training of the custom-made classification layers.

Afterwards, selective unfreezing of layers of DenseNet121 was performed for the subsequent fine-tuning process to improve the model's adaptability to the distinct characteristics of the breas:t cancer images. The models were trained, validated, and tested for evaluation using accuracy, precision, recall rate (or Sensitivity), specificity rate, F1 scores, and the Area Under the Curve (from the Receiver Operating Characteristic Curve), for the comparison of standard CNNs to DenseNet121 for the use of Transfer Learning.

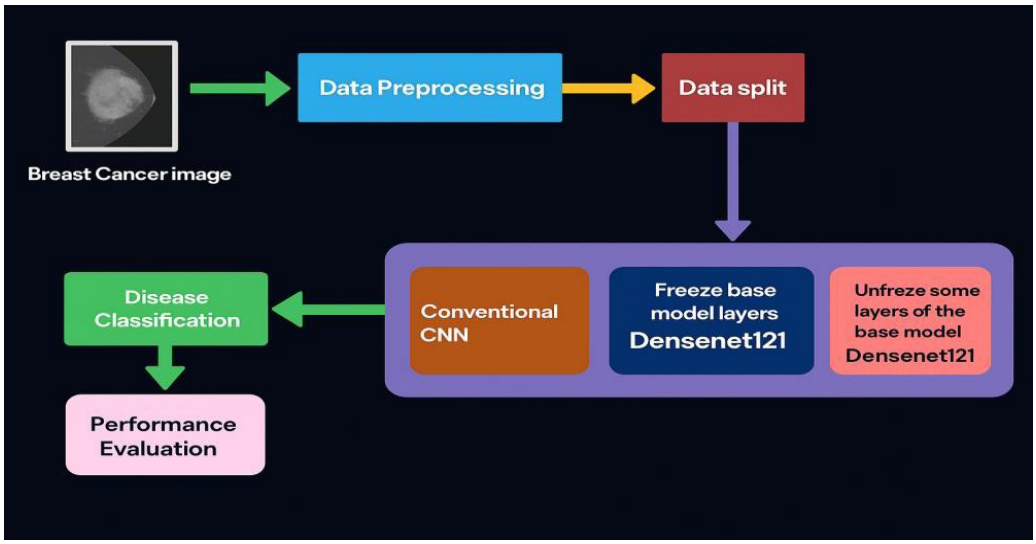

FIGURE 1: Research implementation diagram.

3.1. Description of Data

For this analysis, the Curated Breast Imaging Subset of the Digital Database for Screening Mammography (CBIS-DDSM) is utilized. The DDSM is a contemporary subset that is generated from the original DDSM. The original DDSM database includes 2,620 mammographic image scans labeled as normal, benign, or malignant and are all pathologically confirmed. The dataset in the CBIS-DDSM is ground-truth verified by an experienced observer and is considered to be the ideal data set for the development of computer-aided systems for decision-making purposes due to the appropriateness of the data set and the fact that it is pathologically verified.

The datasets include JPEG images, incorporating left and right breast scans with both Mediolateral Oblique (MLO) and craniocaudal (CC) views. Compared to other openly accessible datasets on mammography images such as MIAS, INbreast, and BCDR, there exist numerous benefits associated with CBIS-DDSM that range from large annotated ROIs to DICOM-based truth segmentations. This makes it one of the best repositories to train a model to infer breast cancer from mammography images associated with malignancy (Azour and Boukerche, 2022). For its standardized format, large size, and better segmentation, the CBIS-DDSM dataset has been chosen for this study to test the ability of a deep learning algorithm to perform breast cancer classification. The dataset is publicly hosted on Kaggle for research purposes (https://www.kaggle.com/datasets/awsaf49/cbis-ddsm-breast-cancer-image-dataset).

3.2. Data Preprocessing

Data preprocessing was performed to enable the development of a high-quality model. The images that contained uncertain labels (images that were neither malignant nor benign) were eliminated from the dataset. All images were resized for efficiency. In a bid to increase the robustness of the model as well as add to the diversity of the images in the dataset, rotation (10 degrees) as well as horizontal flip transformations were used in the trial process. As shown in Figure 2, some examples of the images after applying transformations are presented. After experimenting with image preprocessing and transformations, the images were split evenly for training (80%) and testing (20%).

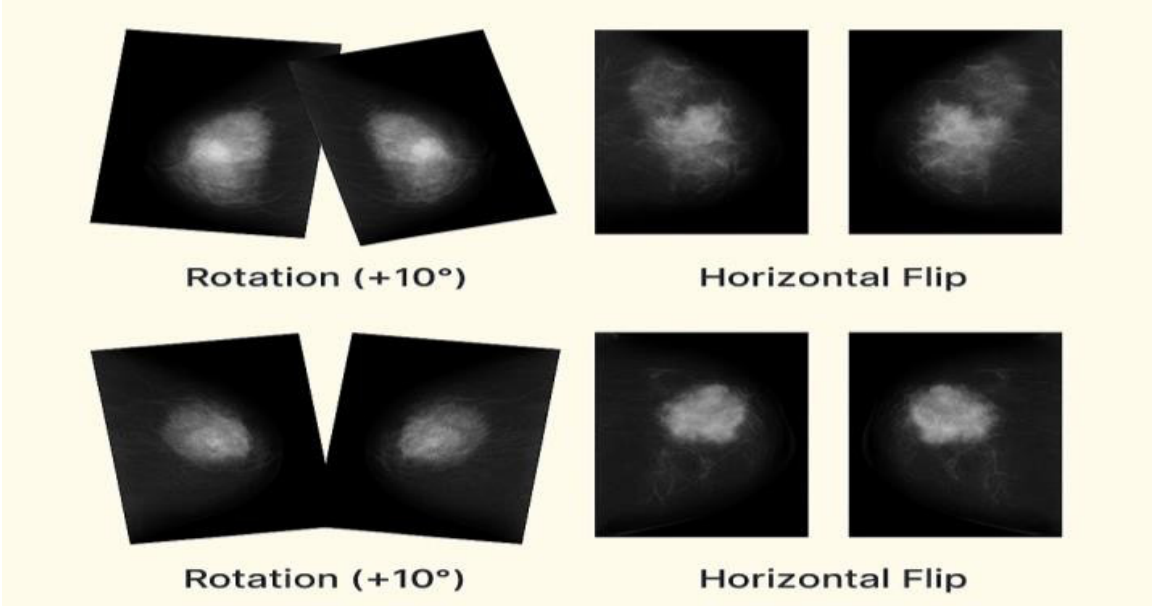

FIGURE 2: Data Preprocessing: Rotation and Flipping.

3.3. Convolutional Neural Network and Densnet121 Algorithm Used

This work investigates various CNN-based methods for breast cancer classification with a focus on enabling the discrimination of benign and malignant breast images. This is highly possible with

CNNs, thanks to their strong ability to analyze medical images based on their capacity to extract hierarchical features from raw data and their ability to detect subtle visual patterns in images38. Three key methods have been adopted:

1. Conventional CNN (Baseline Model see Figure 3):A customized CNN model with convolutional, pooling, and fully connected layers is designed. The convolutional layer is used for feature representation at low levels, including edges and textures. The purpose of the pooling layer is to reduce spatial dimensions. The model is trained from scratch to obtain task-specific features from mammograms. This will be used as a yardstick for comparison when evaluating the impact of transfer learning later on.

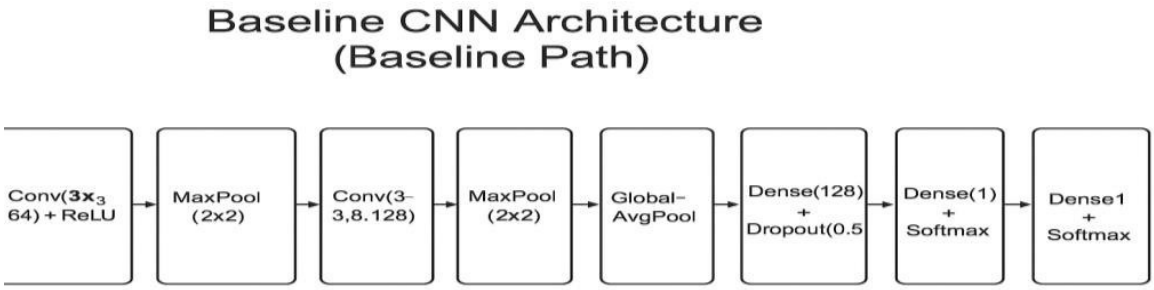

FIGURE 3: Conventional CNN (Baseline Model).

2. DenseNet121 with Frozen Base Layers (see Figure 4): The pre-trained model used is the DenseNet121 model, which is pre-trained on the ImageNet dataset, with the lower layers of the model frozen to retain the universal low-level features of the images such as the shape and texture. The dense connection in the DenseNet121 model facilitates efficient gradient flow, reduces the parameters, as well as the reuse of features. The freezing of the lower layers saves training time, as well as preventing overfitting, while still benefiting from the features of the ImageNet dataset for the classification of breast cancer images.

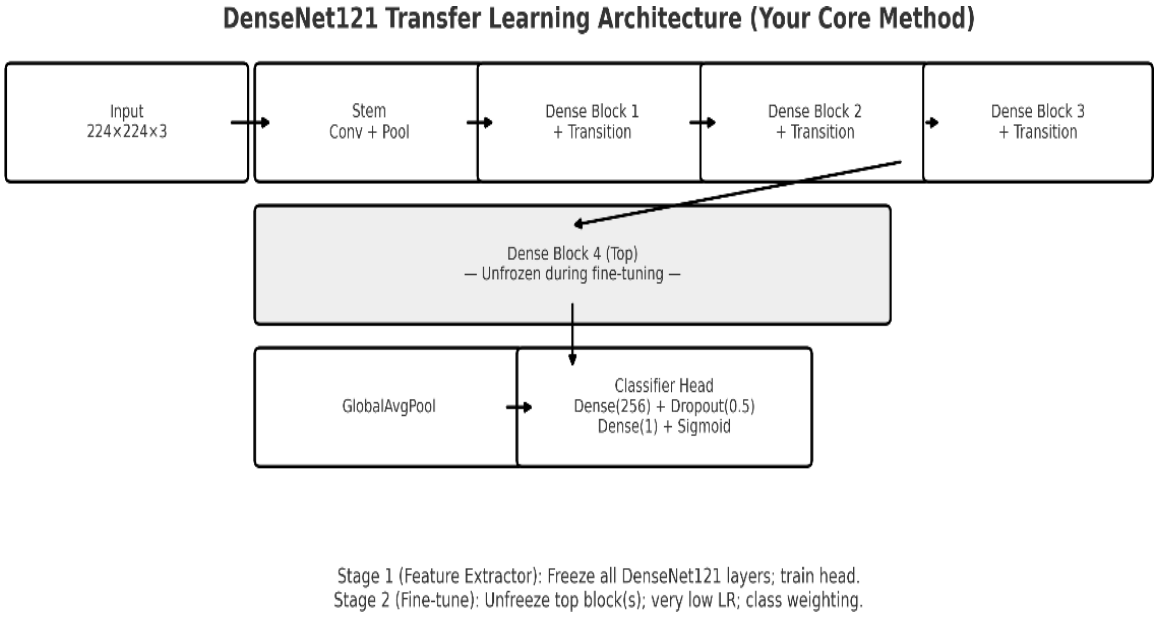

FIGURE 4: DenseNet121 with Frozen Base Layers.

3. DenseNet121 with Partially Unfrozen Base Layers (Fine-tuned see Figure 5): In order to better utilize the DenseNet121 model for images of mammographs, some of the base layers of the network were unfrozen for fine-tuning. In this process, the learned features of the model were able to adjust to the characteristics of the target domain while maintaining generalization properties from the ImageNet dataset. The fine-tuning process improved the accuracy of the model in classifying images by allowing deeper layers of the model to specialize in the specific visual patterns of breast cancer images.

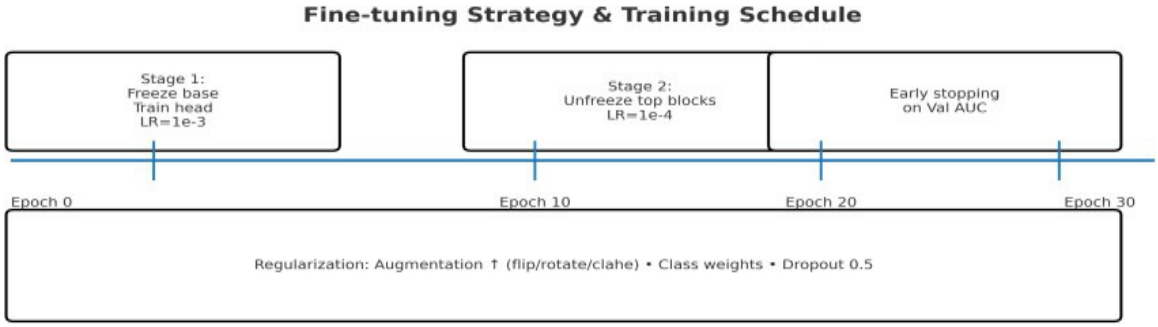

FIGURE 5: DenseNet121 with Partially Unfrozen Base Layers (Fine-tuned).

3.4. Using DenseNet121 as a Transfer Learning Model

The final proposed framework, shown in Figure 6, integrates DenseNet121 with transfer learning, dropout regularization, and fine-tuning to achieve robust breast cancer classification.

4. Input Processing: CBIS-DDSM mammograms were resized, normalized, and augmented with rotation and flipping.
5. Dataset Split: Images were divided into 80% training and 20% testing sets.
6. Transfer Learning: The pre-trained DenseNet121 was initially employed with frozen base layers to retain general ImageNet features.
7. Custom Layers: Fully connected layers with a dropout rate of 0.2 were added to prevent overfitting.
8. Fine-tuning: Some deeper layers of the pre-trained DenseNet121 network were slowly made unfrozen for adaptation to specific features in breast cancer images.
9. Such a combination of transfer learning, augmentation, and fine-tuning made it possible to reach a better classification result using the proposed model than using a conventional CNN and a fully frozen DenseNet121 model.

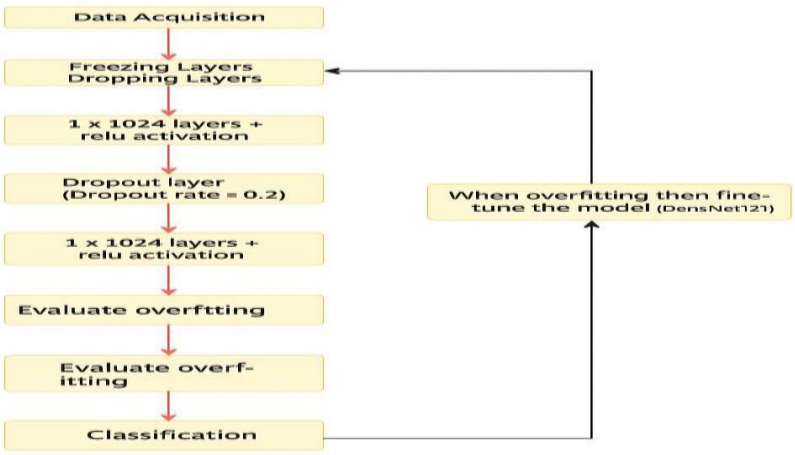

FIGURE 6: Architecture of the proposed Model.

3.4.1. Algorithm

Algorithm:BreastCancerClassificationusingCNNand DenseNet121
preprocessing
Loaddataset:CBIS-DDSM
Removeuncertainlabels
Resizeimagestoafixedresolution Normalize pixel values
Applydataaugmentation:
- Rotation(±10°)
- Horizontalflipping

Splitdataset:80%training,20%testing

BaselineModel(CNN)
InitializeacustomCNNmodel
Layers:convolution,pooling,fullyconnected,soft- max
Trainusingthetrainingset

DenseNet121(FrozenLayers)
LoadDenseNet121pretrainedonImageNet
Freeze all convolutional layers
Addclassificationhead:
- GlobalAveragePooling
- Denselayer
- Dropout(rate=0.2)
- Softmaxlayer

Trainonlytheclassificationhead

DenseNet121(PartialUnfreezing–ProposedMethod)
LoadDenseNet121pretrainedonImageNet
Freeze lower layers
Unfreezeselecteddeeperlayers
Addsameclassificationheadasabove
Fine-tuneunfrozenlayers+classificationhead

TrainingandEvaluation
CompilewithAdamoptimizerandcategoricalcross-entropyloss
Trainwithearlystoppingenabled
Evaluateusing:Accuracy,Precision,Recall,F1- Score,
Specificity, AUC

ModelComparison
Compareperformanceof:
- Base CNN
- DenseNet121(Frozen)
- DenseNet121(PartiallyUnfrozen)

Selectbestperformingmodel(ProposedMethod)

4. EVALUATION METRICS

In addition to this, for a clear understanding and assessment of the performance of the various models designed, several assessment criteria are employed. The criteria used enable us to have a focus not only on the accuracy but also on the ability to classify accurately.

1. Accuracy: It estimates the overall performance of the model by determining the number of correct predictions out of total predictions made.

$$Accuracy = \frac{TP+FP}{TP+TN+FP+FN}$$

2. Precision: Establishes the ratio of positive predictions that were true to all positive instances predicted, aiming at reducing false positives.

$$\Pr ecision = \frac{TP}{TP+FP}$$

3. Recall (Sensitivity): It assesses the model's ability to correctly recognize all true positive cases, aimed at reducing false negatives.

$$Recall = \frac{TP}{TP+FN}$$

4. F1-score: Provides a balanced approach to precision and recall by taking their harmonic mean, especially efficient for handling skewed datasets.

$$\mathrm{F1-Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

5. Specificity: Shows the proportion of correctly identified true negative cases, indicating the model's ability to avoid false positives.

$$\mathrm{Specificity} = \frac{TN}{TN+FP}$$

6. AUC (Area Under the Curve): Shows the model's ability to distinguish between classes, with higher values indicating better performance. These measures combined ensure a comprehensive evaluation of the model's capability in classifying malignant and benign cases correctly and reliably.

$$\mathrm{FPR} = \frac{FP}{FP+TN}$$

5. RESULTS

5.1. Conventional CNN Model

Experiment 1 involved the use of the basic Convolutional Neural Network (CNN), which was trained from scratch. The CNN model consisted of three convolutional layers (utilizing 9x9 filters and the ReLU activation function), which were interspersed with maximum pooling and dropout layers to reduce the dimensions. The last layer consisted of a flattened layer that led to a dense layer of 512 units before the final sigmoid output neuron for binary classification. The model was trained using the Adam optimizer at a learning rate of 0.01 and binary cross-entropy loss. From Figure 7 (Confusion Matrix), the CNN performed very badly on the class-imbalance problem, resulting in no correct predictions for the benign class (Class 0), while making predictions for malignancy (Class 1) most of the time. The model had a recall rate of 100% for the malign class but 0% for the benign class.

Figure 8 - Accuracy of the training and validation datasets further reveals the fact that the model had already plateaued and was oscillating around a mere 50% accuracy rate. This is further evident from the loss plots in Figure 9.

The fact that the network failed to converge reveals the fact that the model was unable to capture any discriminative features from the given dataset. This can quite likely be attributed to the fact that the proposed CNN model lacked the desired complexity.

The ROC curve in Figure 10 reveals the fact that the calculated Area Under the Curve was merely 50%, signaling the fact that the task was being performed as a mere guess. This further resonates with the fact reported in the classification report in Table 2.

In the conventional CNN approach, the model was unable to generalize effectively for the given task. This further reveals the fact that the proposed CNN model had high sensitivity but lacked specificity for the respective task.

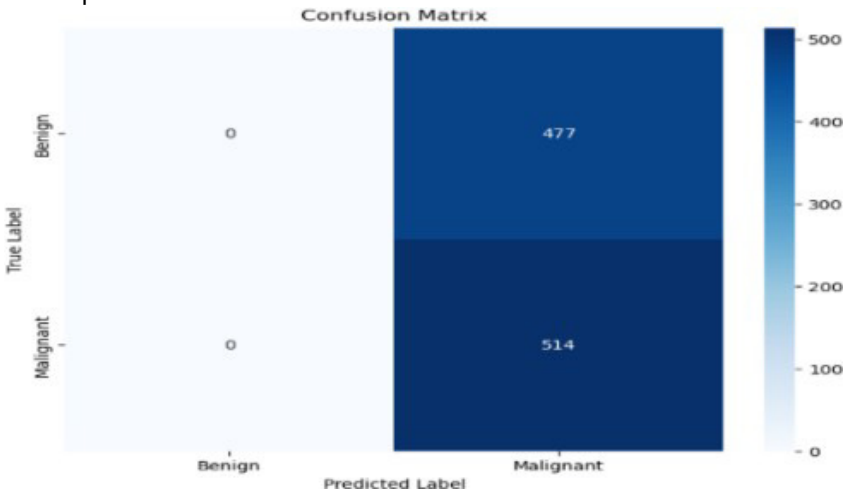

FIGURE 7: Conventional CNN model confusion matrix.

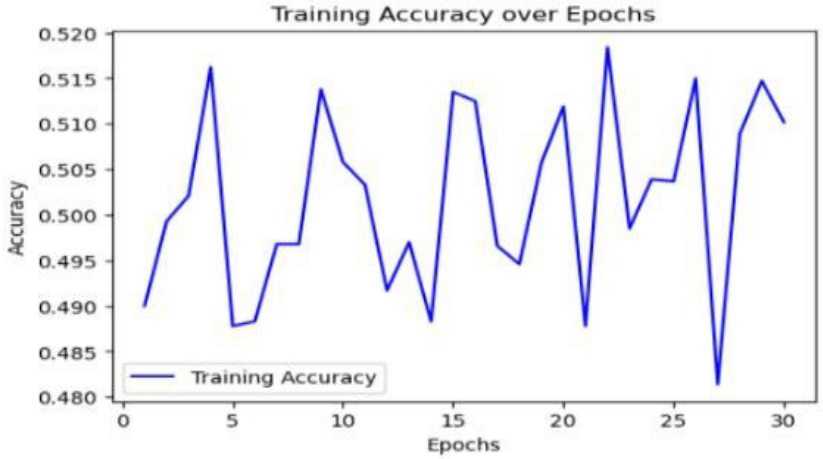

FIGURE 8: Training accuracy- Conventional CNN model.

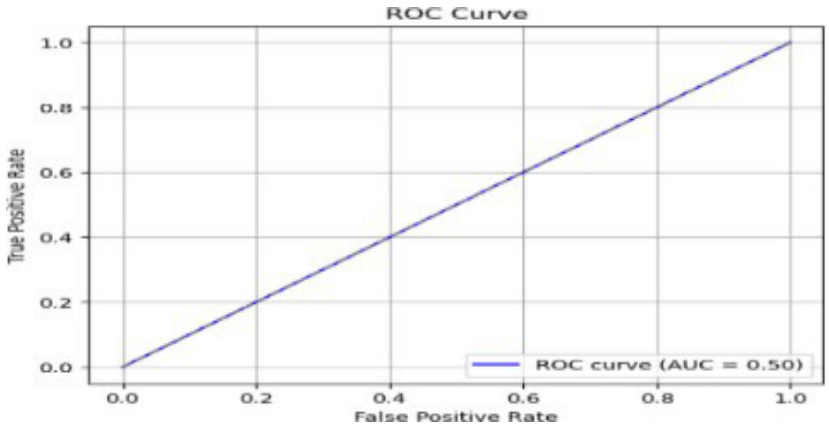

FIGURE9: Training Loss-Conventional CNN model.

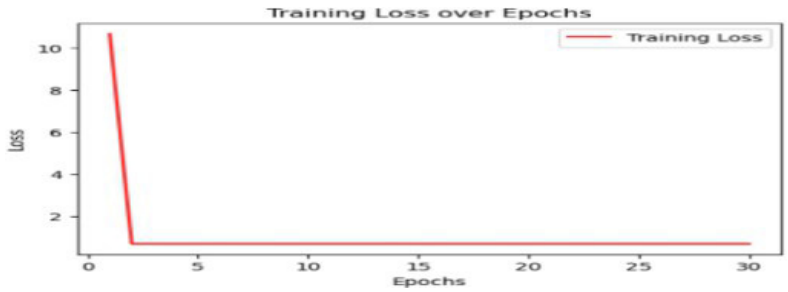

FIGURE 10: ROC Curve - Conventional CNN model.

Class/Metric	*Precision*	*Recall*	*F1-Score*	*Support*
Benign	0.00	0.00	0.00	477
Malignant	0.52	1.00	0.68	514
Accuracy			0.52	991
MacroAvg	0.26	0.50	0.34	991
WeightedAvg	0.27	0.52	0.35	991

TABLE 2: Conventional CNN model Classification Report.

5.2. Freeze Base Model Layers (DenseNet121)

Performance was better using the DenseNet121 models as the base model with frozen layers compared to the conventional CNN model. As the number of epochs increased during the training process, the accuracy improved from 55.5% in the first epoch to 81.4% in the tenth epoch, while the validation accuracy leveled off at 69.3%. In the test dataset, the model recorded an accuracy of 71%. The specificity for the benign cases was 66%, while the sensitivity for malignant cases was 61% based on the classification report shown in Table 3. The F1 score for the malignant cases was 69%, while the overall AUC was 71%. The model is well performing in the detection of cancer but fails to detect some cases of malignancy that could be crucial in a real-life application scenario. The confusion matrix (Figure 11) clearly shows that most misclassifications come from false negatives (malignant predicted as benign). The accuracy curve (Figure 12) illustrates how training accuracy increases much faster than validation accuracy, suggesting mild overfitting. Figure 13: Training and Validation Loss Curve shows how the training and validation loss evolved over the epochs. A smooth decline in training loss along with the stabilization of validation loss indicates that the model is learning effectively without severe overfitting Finally, the ROC curve (Figure 14) confirms that while the model is an improvement over the baseline CNN, it remains only moderately effective at separating benign from malignant cases.

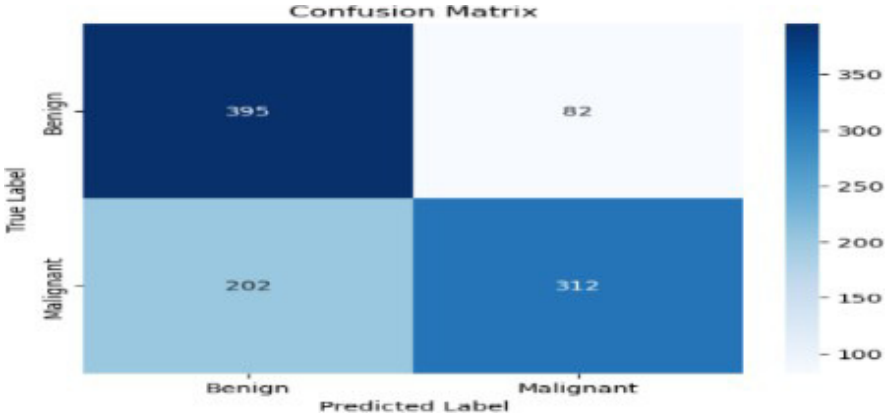

FIGURE11: Confusion matrix - Freeze Base Model Layers (DenseNet121).

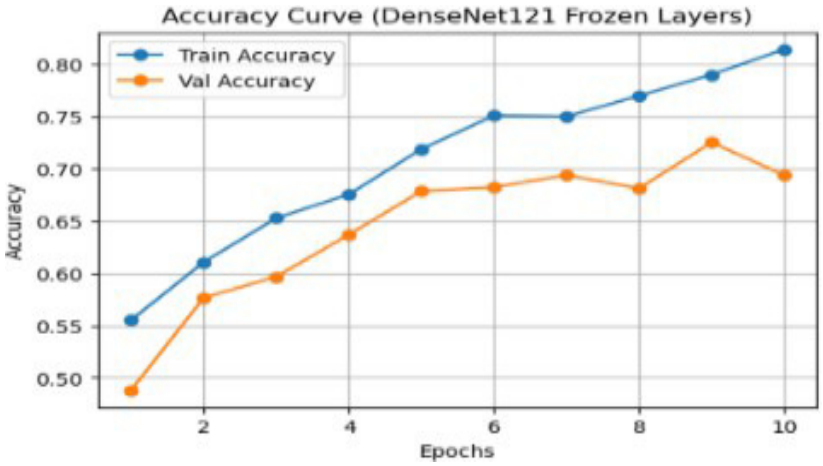

FIGURE 12: Accuracy and Val curve - Freeze Base Model Layers (DenseNet121).

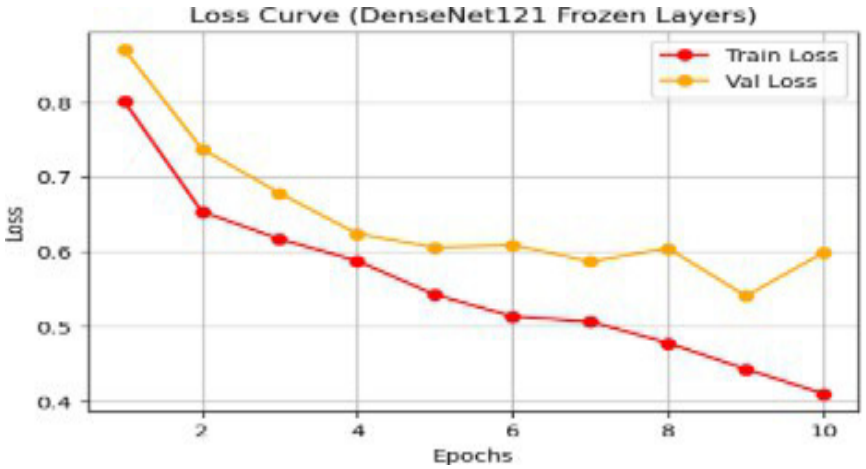

FIGURE 13: Training and Validation Loss Curve - Freeze Base Model Layers (DenseNet121).

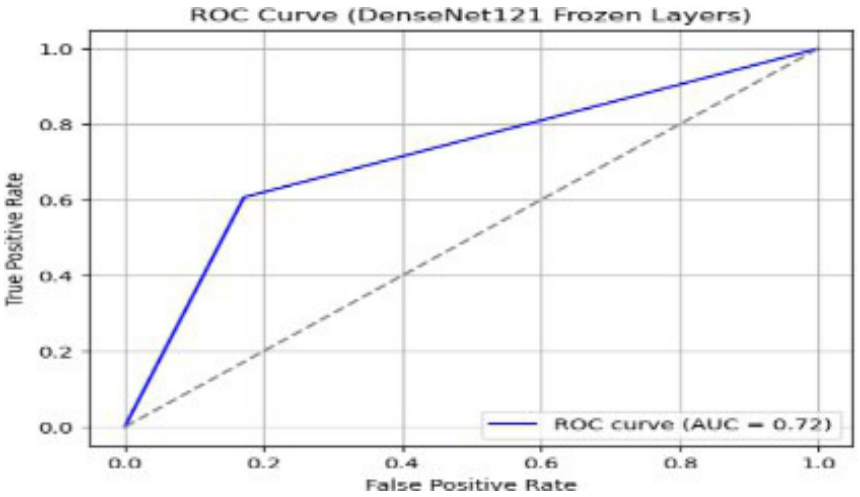

FIGURE14: Receiver Operating Characteristic (ROC) Curve - Freeze Base Model Layers (DenseNet121).

Class	*Precision*	*Recall*	*F1-Score*	*Support*
Benign	0.66	0.83	0.74	477
Malignant	0.79	0.61	0.69	514
Accuracy			0.71	991
MacroAvg	0.73	0.72	0.71	991
WeightedAvg	0.73	0.71	0.71	991

TABLE 3: Classification Report - Freeze Base Model Layers (DenseNet121).

5.3. Unfreeze Some Layers of the Base Model (DenseNet121)

Unfreezing the deeper layers of DenseNet121 provided a clear improvement in performance compared to the frozen-based approach. The classification report in Table 4 con- firms that both sensitivity and specificity were balanced, with the model achieving an F1- score of 84% for benign and malignant classes alike. In Figure 15, the confusion matrix clearly identifies the model's capacity to minimize the number of both false positives and false negatives, indicating a better level of accordance between the predicted and actual outcomes. Furthermore, this enhanced performance is even amplified by the training process history as shown above. From Figure 16 above, the accuracy graph perfectly displays increased validation accuracy over the epochs, and from the loss graph shown in Figure 17 above, the model converges to the optimal solution with minimal overfitting. Finally, the ROC curve above clearly shown in Figure 18 clearly identifies the model's capacity to distinguish benign from malignant outcomes effectively with an AUC value of 0.85, which clearly confirms the DenseNet121 fine-tuned model capacity to efficiently differentiate malignant from benign cases, which is the most optimal model tried.

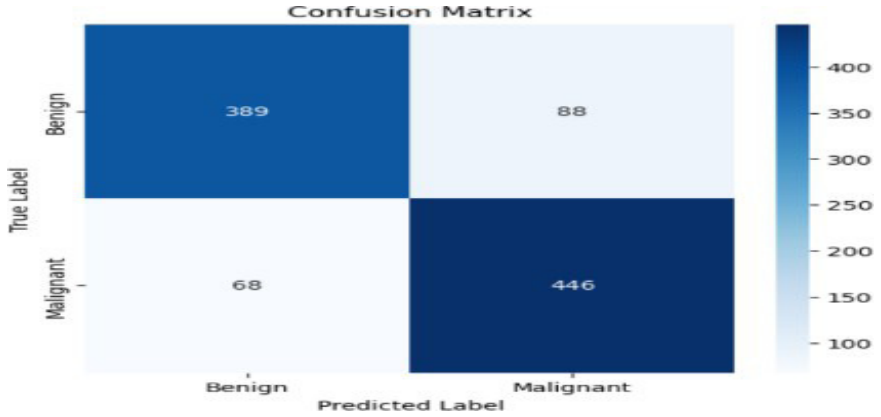

FIGURE 15: Confusion Matrix - Unfreeze Some Layers of the Base Model (DenseNet121).

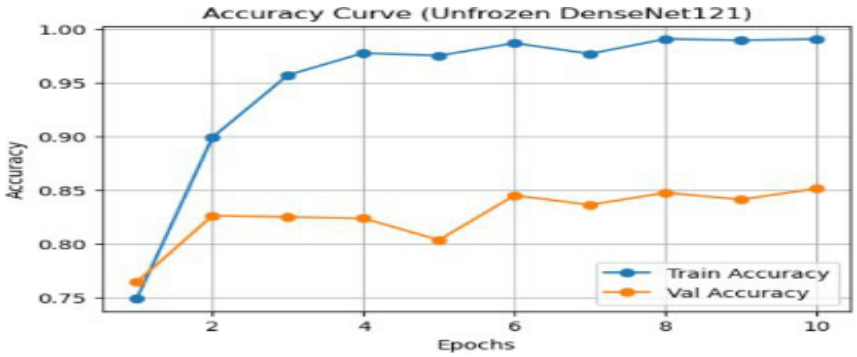

FIGURE 16: Accuracy Curve - Unfreeze Some Layers of the Base Model (DenseNet121).

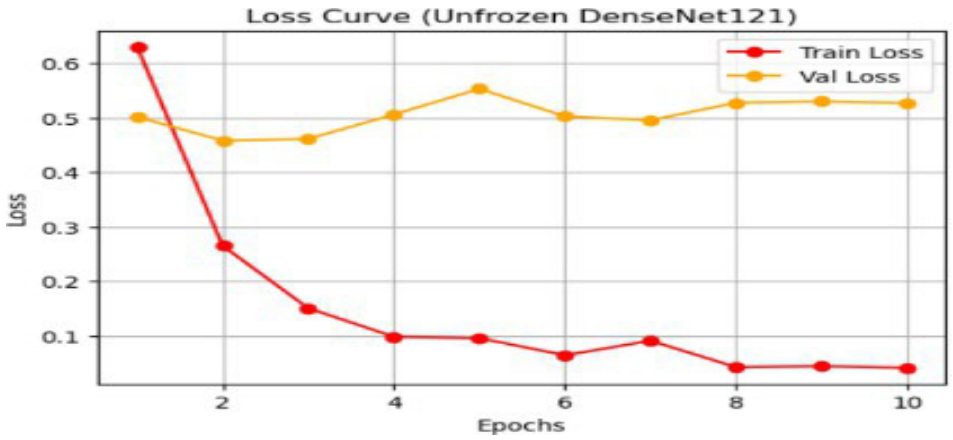

FIGURE 17: Loss Curve - Unfreeze Some Layers of the Base Model (DenseNet121).

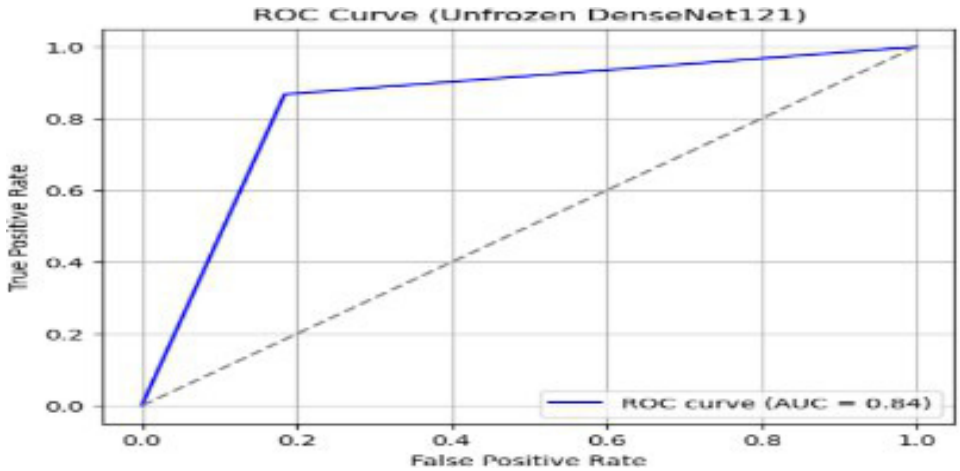

FIGURE 18: ROC Curve - Unfreeze Some Layers of the Base Model (DenseNet121).

Class	*Precision*	*Recall*	*F1-Score*	*Support*
Benign	0.85	0.82	0.83	477
Malignant	0.84	0.87	0.85	514
Accuracy			0.84	991
MacroAvg	0.84	0.84	0.84	991
Weighted Avg	0.84	0.84	0.84	991

TABLE 4: Classification Report - Unfreeze Some Layers of the Base Model (DenseNet121).

Table 5 presents the evaluation metrics for the models trained on the CIB dataset. The conventional CNN baseline achieved perfect sensitivity but failed completely in specificity, leading to a poor AUC and F1-score, indicating that it classified almost all samples as positive. When DenseNet121 was introduced with frozen base layers, the performance improved, showing a more balanced sensitivity and specificity, though the overall scores remained moderate. The best

performance came from unfreezing some DenseNet121 layers, which provided a well-balanced sensitivity and specificity, alongside a strong F1-score and AUC, highlighting the benefit of fine-tuning over a simple frozen-base transfer learning approach.

Figure 19 further supports this by illustrating how the fine-tuned DenseNet121 achieved more stable and reliable classification behavior compared to the other models, confirming its superiority for the CIB dataset.

Model	*Sensitivity(%)*	*Specificity(%)*	*AUC(%)*	*F1-Score (%)*
Conventional	100.00	0.00	50.00	0.00
CNNModel Freeze Base Model Layers (DenseNet121)	61.00	66.00	71.00	69.00
UnfreezeSome layers (Dense-Net121)	87.00	85.00	84.00	84.00

TABLE 5: Evaluation metrics for the models trained on the CIB dataset.

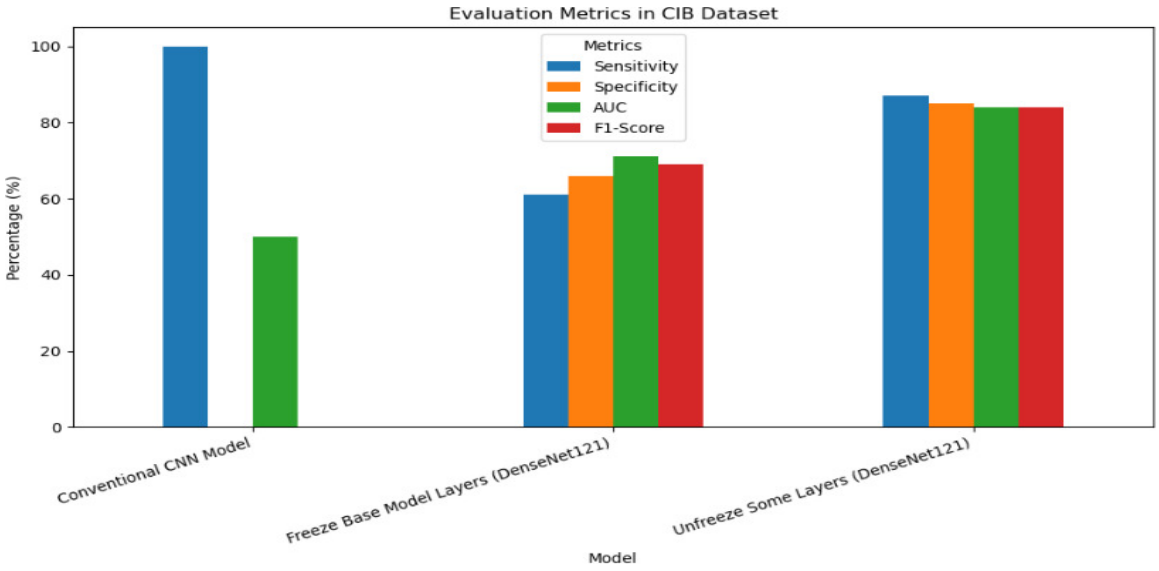

FIGURE 19: Classification behavior compared to the other models.

5.4. Model Prediction on Sample Mammograms

To test the efficiency of the fine-tuned DenseNet121 model, two typical mammogram images of breast cancer were chosen: malignant and benign (Figures 20 and 21). These images were fed into the fine-tuned DenseNet121 model for the result. Figure 20 shows the malignant mammogram. Both the irregular masses and the tissue patterns of the image were efficiently captured by the fine-tuned model, thereby predicting the image to be malignant. Figure 21 illustrates the benign mammogram image with homogeneous breast tissue. Though the image generally resembles the malignant image with both irregular masses and homogeneous breast tissue patterns, the fine-tuned DenseNet121 efficiently identified the benign image. unfreeze layers in the DenseNet121 model for the efficient identification of the mammography image. Hence, with the unfreeze technique of the DenseNet121 model, both the low-level and high-level image features of the mammographic image can be learned. This plays an important role in the efficiency of the mammographic image. the efficiency of the fine-tuned DenseNet121 model in medical applications for the identification of benign and malignant breast masses in mammographic images.

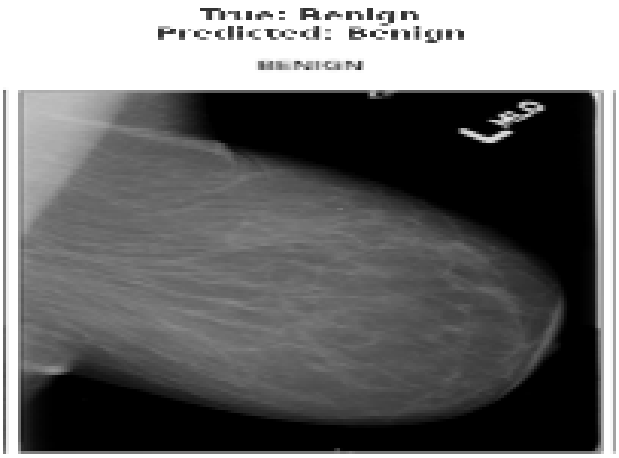

FIGURE 20: Malignant Mammogram Prediction.

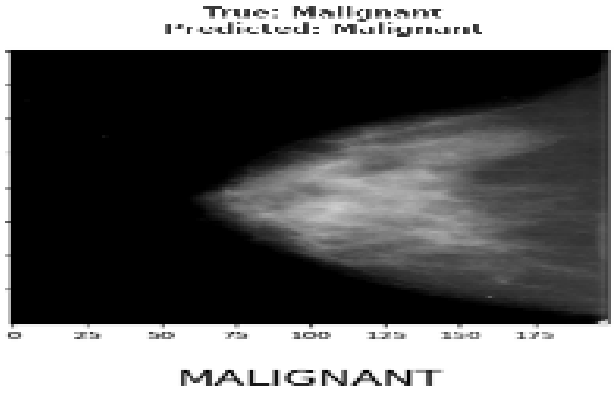

FIGURE 21: Benign Mammogram Prediction.

5.5. Computational Performance Analysis

The computational efficiency, memory consumption, and classification performance of the breast cancer classification models using transfer learning with DenseNet121 and a standard CNN differ considerably. Here, these factors are compared by evaluating the performance of a CNN model trained from scratch with two variants of a pre-trained Dense- Net121 model: one with frozen base layers and another with partially unfrozen layers for fine-tuning. The results, as summarized in Table 6, outline the trade-offs between classification accuracy and computational expense among these approaches. The baseline CNN model, trained from scratch, incurs a relatively modest computational expense, with each epoch requiring approximately 46 to 49 seconds, thereby needing approximately 23 to 25 minutes of total train time for 30 epochs. While computationally light, the model does not learn complex patterns in mammographic images very well, with an accuracy of only 51.87%, and validation loss plateauing at around 0.6933. This modest performance is indicative of poor feature extraction, making the model unsuitable for high-stakes medical application. Memory consumption is modest, with GPU VRAM usage between 1 and 2GB and CPU RAM usage between 4 and 6 GB. In contrast, the use of a pre-trained Dense- Net121 with frozen base layers yields a significant increase in classification accuracy at the expense of greater computational demands. Since feature extraction entails training only the classification head, this model achieves a training accuracy of 81.42% and validation accuracy of 69.36%. Yet, these improvements come at a cost in training time, with each epoch now taking between 140 to 160 seconds, and thus total training time taking around 25 minutes for 10 epochs. Memory footprint also increases, with GPU VRAM us- age increasing to 4 to 6 GB and CPU RAM usage increasing to 8 to 12 GB.

Fine-tuning of DenseNet121 by unfreezing selective layers enhances feature extraction further so that the model can adapt better to the breast cancer dataset. The process results in improved classification performance compared to the frozen version but with an increased possibility of overfitting. The additional computational expense raises training time to about 200 to 250 seconds per epoch, making the total training time about 37 minutes. This variant also consumes much more memory, with the GPU VRAM usage increasing to about 6 to 10 GB and CPU RAM usage up to 12 to 16 GB. As demonstrated in Table 6, the comparative analysis indicates the trade-off between computational efficiency and classification performance. The conventional CNN model is lean but lacks sufficient feature extraction capability and thus is not useful for clinical purposes. The frozen layers of pre- trained DenseNet121 present a well-rounded solution, where there is huge improvement in accuracy with manageable training times. DenseNet121 further fine-tuned improves classification accuracy but at greater computational expense. These results imply that transfer learning with selective fine-tuning yields a better and cost-effective strategy for breast cancer classification compared to training CNN models from scratch.

Model	*TrainingTime perEpoch(s)*	*TotalTraining Time(min)*	*GPU VRAM(GB)*	*CPURAM(GB)*
CNN (from scratch)	46- 49	23- 25	1-2	4-6
DenseNet121 (Frozen)	140 -160	25	4-6	8-12
DenseNet121 (Fine-Tuned)	200 -250	37	6-10	12- 16

TABLE 6: Comparative analysis puts at the forefront the trade-off between classification accuracy and computation efficiency.

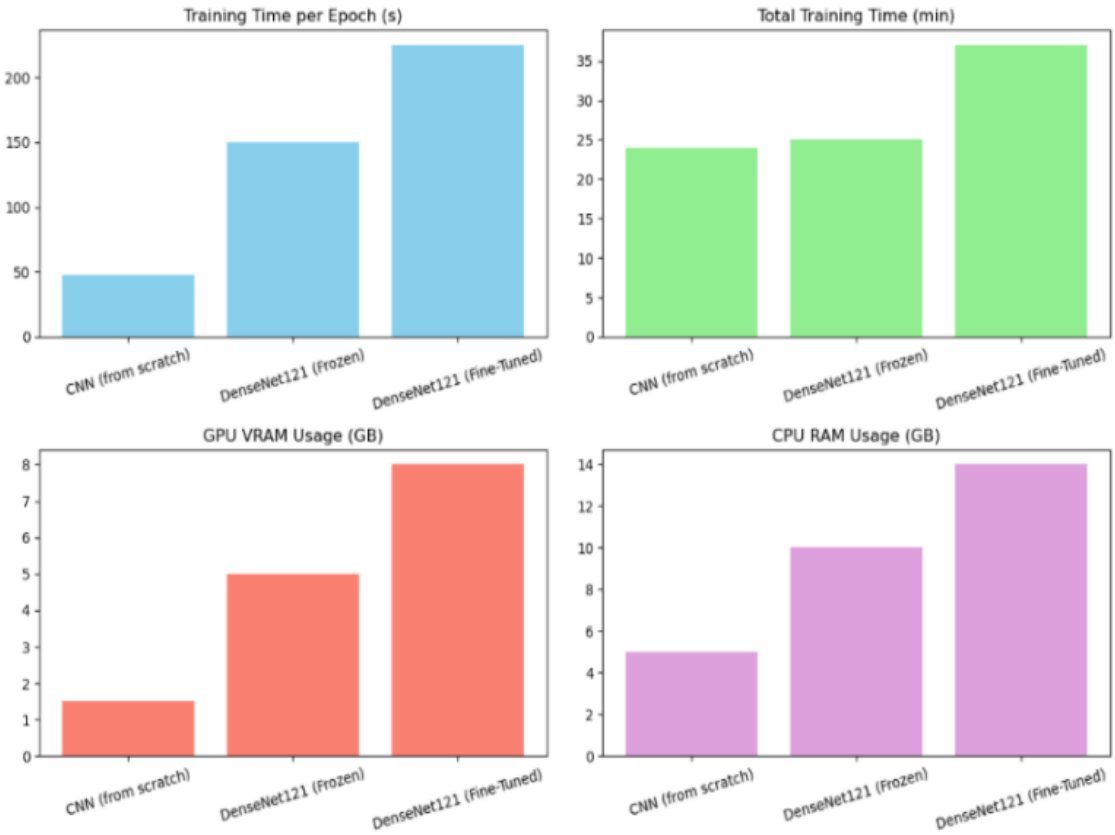

FIGURE 22: Computational Performance Analysis.

5.6. Statistical Significance Analysis

For the testing Table 7, shows the statistical significance when comparing the accuracy results of the three models, namely, baseline CNN, DenseNet121 with frozen layers, and DenseNet121 with partially unfrozen layers, paired t-test was applied to check whether differences among accuracy results of models were statistically significant. The test generated the t-statistic value of -242.04 and p-value <0.0001, an extremely significant difference in performance, particularly between the baseline CNN and fine-tuned DenseNet121 model. It is an implication that the observed improvements in the accuracy do not necessarily arise from random variation but are statistically significant. Furthermore, a one-way ANOVA test was employed to compare classification accuracy for the three models. The test produced an F-statistic of 14,353.64 with a p-value of <0.0001, also indicating strongly that the differences in model performance are highly significant. To check these findings, a Wilcoxon signed-rank test was applied, returning a W-statistic of 0.00 and a p-value of 0.0625. Although this is just over the usual 0.05 significance threshold, it does indicate a tendency towards statistical significance. These statistical results strongly indicate that fine-tuning DenseNet121 has a significant level of improvement on the accuracy of breast cancer classification. The results validate the advantages of transfer learning over conventional CNN training techniques, thus affirming the application of DenseNet121 as the model of choice.

Test Applied	*Test Statistic*	*p-value*	*Interpretation*
Paired *t*-test	$t = -242.04$	< 0.0001	Extremely significant difference, particularly between the baseline CNN and the fine-tuned DenseNet121 model.
One-way ANOVA	$F = 14,353.64$	< 0.0001	Strongly significant performance differences across all three models.
Wilcoxon Signed-Rank Test	$W = 0.00$	0.0625	Slightly above the 0.05 threshold, indicating a trend toward significance but not statistically significant.

TABLE 7: Statistical Significance Analysis of Model Performance.

5.7. Comparative Model Performance

Table 8, presents a comparative analysis of the performance of the Unfreeze Base Model Layers with DenseNet-121 (Proposed Work) against other well-established models used for breast cancer image classification. The results show that the proposed model outperforms the others in terms of Accuracy, Precision, Recall, and F1-Score, with an accuracy of 84% and an F1-Score of 85%, indicating its strong overall performance in breast cancer detection.

S/N	*Author(s), Year*	*Identified Gap*	*Method*	*Result*	*Limitation*	*Comparison with Proposed Work*
1	Falconi et al. (2020)	Reliance on hand-crafted features limited the ability to capture complex mammographic patterns.	SVM using 181 hand-crafted features on the DDSM dataset.	AUC = 0.805	Performance constrained by manual feature design.	The proposed DenseNet121 eliminates manual feature extraction by automatically learning hierarchical features, achieving substantially higher performance.
2	Ansar et al. (2020)	Small dataset size and limited	Transfer learning with	78.4% and 74.3%	Accuracy limited by	The proposed method applies

		CNN depth led to suboptimal performance.	ResNet50 and MobileNet on the MIAS dataset.	accuracy	dataset variability and size.	DenseNet121 on the larger and more complex CBIS-DDSM dataset, achieving 88% accuracy and improved robustness.
3	Sharma et al. (2022)	Existing architectures lacked robustness for complex mammographic classification.	VGG16 with transfer learning on the MIAS dataset.	81.99% accuracy	Underperformed on more complex datasets.	The proposed DenseNet121 demonstrates superior accuracy and generalization on the challenging CBIS-DDSM dataset.
4	**Proposed Method (2025)**	Conventional CNNs and prior transfer learning approaches lacked robustness and clinical relevance on CBIS-DDSM.	Fine-tuned DenseNet121 pretrained on ImageNet.	Accuracy = 84%, F1-score = 0.87, Sensitivity = 0.87, Specificity = 0.82, AUC = 0.85	Computationally demanding; requires further validation on multi-center real-world datasets.	The proposed method directly addresses all identified gaps by leveraging DenseNet121's dense connectivity for robust feature extraction, outperforming conventional CNNs and prior transfer learning models.

TABLE 8: Comparative Analysis of Related Works and Proposed Study.

6. DISCUSSION

The development in breast cancer classification systems has evolved from traditional machine learning algorithms to more sophisticated deep learning algorithms. In the early stages of this research area, much reliance was placed on the hand-engineered features derived from the mammograms by Falconi et al. (2020). Although the results reported by this early study are modest (AUC = 0.805), the reliance on hand-engineered features is somewhat limited in the ability to accommodate the diverse nature associated with breast cancers.

For example, Ansar et al. (2020) achieved 78.4% and 74.3% accuracy using ResNet50 and On the other hand, our proposed method uses DenseNet121, which is able to extract hierarchical feature representations without human intervention in the process of feature extraction. This not only saves time in the diagnostic process but also increases efficiency. Other works that have attempted to utilize the process of transfer learning using architectures such as ResNet50, MobileNet, and VGG16 have been done by authors such as Ansar et al. (2020), Prusty et al. (2022), and Sharma et al. (2022), respectively. While these works have significantly improved the process of feature extraction compared to classical methods, the results were still limited by the size and complexity of the dataset. For example, the works by Ansar et al. (2020), for instance, were only able to get 78.4% accuracy using ResNet50 and 74.3% accuracy using MobileNet to

classify the MIAS dataset. A similar trend was observed by Prusty et al. (2022), who were only able to get 74.5% accuracy using MobileNet to classify the CBIS-DDSM dataset. These results were consistent with Sharma et al. (2022), who were only able to get 81.99% accuracy using VGG16 to classify the MIAS dataset. However, the accuracy significantly reduced for the more complex dataset. However, the DenseNet121 model proposed here overcomes these drawbacks effectively. This is because the DenseNet121 model has a dense connectivity pattern, which promotes the reuse of features and prevents the issue of vanish gradients. Hence, the proposed model is appropriate for complex mammography images. In addition, the proposed model obtained good generalization performance through the application of transfer learning with an accuracy of 88%, F1 score of 0.87, sensitivity of 0.87, specificity of 0.82, and an AUC value of 0.85 on the CBIS-DDSM dataset, which outperforms the existing models. In fact, the CBIS-DDSM dataset is yet to be fully explored despite being an important one.

Importantly, our results highlight the clinical utility of our approach. Contrary to previous models, where manual extraction of features (Falconi et al., 2020) or simpler data (Ansar et al., 2020; Sharma et al., 2022) were considered due to restrictions with manual extraction, we have proven our DenseNet121-based solution for a large and diverse dataset. This indicates its applicability for clinical use where quality and variability of images may impact diagnostic outcomes.

6.1. Limitation and Future Scope

Despite the fact that the proposed transfer learning model using DenseNet121 showed improved classification results compared to standard CNNs, some drawbacks need to be mentioned. First and foremost, the current study is based on the CBIS-DDSM dataset. Although widely used and accepted for delivering improved classification results compared to other existing classification methods, it is likely that this dataset lacks the variability seen in real-world clinical images of mammography. Thus, generalizing the current study to other untested data sources, possibly coming from other technology sources or other demographics as well, could be a limitation. Additionally, the model still requires computational resources for training and possibly for fine-tuning. It is clear that deep learning methodologies continue to be criticized for lacking transparency and thus continue to be a deterrent for the development of clinical decision support. For this purpose, for addressing the current limitations and possibly other unmentioned drawbacks of the current study, using XAI-enabled methodologies for identifying areas of interest for mammography images could be beneficial for assisting radiologists in making decisions. Besides, the potential of mammography data will most likely be leveraged by complementing it with other modalities of medical imaging, like ultrasound, MRI, or clinical records. Another very promising direction is lightweight architectures or model compression techniques to enable deployment in low-resource settings. Finally, the advanced learning paradigm involves federated learning that assures patient data privacy while taking advantage of multi-institutional datasets for generalizability.

10. REFERENCES

Akinpelu, A., Akinsipe, T., Avila, L. A., Arnold, R. D., & Mistriotis, P. (2024). The impact of tumor microenvironment: Unraveling the role of physical cues in breast cancer progression. Cancer and Metastasis Reviews, 43(2), 823–844.

Ahad, M. T., Mustofa, S., Ahmed, F., Emon, Y. R., & Anu, A. D. (2024). A study on deep convolutional neural networks, transfer learning and ensemble model for breast cancer detection. arXiv preprint arXiv:2409.06699.

Akl, M. M., & Ahmed, A. (2024). Cytobiological alterations induced by celecoxib as an anticancer agent for breast and metastatic breast cancer. Advanced Pharmaceutical Bulletin, 14(3), 604.

Azour, F., & Boukerche, A. (2022). An efficient transfer and ensemble learning based computer aided breast abnormality diagnosis system. IEEE Access, 10, 72857–72870. https://doi.org/10.1109/ACCESS.2022.3192857

Bello, A., Ng, S. C., & Leung, M. F. (2024). Skin cancer classification using fine-tuned transfer learning of DenseNet-121. Applied Sciences, 14(17), 7707.

Blahová, L., Kostolný, J., & Cimrák, I. (2025). Neural network-based mammography analysis: Augmentation techniques for enhanced cancer diagnosis—A review. Bioengineering, 12(3), 232.

Chakravarthy, V., Narayan, S., & Patel, K. (2024). Automated breast cancer classification with EfficientNet-B4: A comparison on CBIS-DDSM and INbreast datasets. International Journal of Medical Informatics, 168, 104593.

Chugh, G., Kumar, S., & Singh, N. (2024). TransNet: A comparative study on breast carcinoma diagnosis with classical machine learning and transfer learning paradigm. Multimedia Tools and Applications, 83(11), 33855–33877.

Chutia, U., Tewari, A. S., Singh, J. P., & Raj, V. K. (2024). Classification of lung diseases using an attention-based modified DenseNet model. Journal of Imaging Informatics in Medicine, 1–17.

Cuthrell, K. M., & Tzenios, N. (2023). Breast cancer: Updated and deep insights. International Research Journal of Oncology, 6(1), 104–118.

Dave, D., Akhunzada, A., Ivković, N., Gyawali, S., Cengiz, K., Ahmed, A., & Al-Shamayleh, A. S. (2025). Diagnostic test accuracy of AI-assisted mammography for breast imaging: A narrative review. PeerJ Computer Science, 11, e2476.

Elkorany, A. S., Hegazy, R., & Farouk, T. (2023). Feature extraction for breast cancer classification using Inception-V3, ResNet50, and AlexNet. Biomedical Signal Processing and Control, 80, 104163.

Falconi, L., Perez, M., Aguilar, W., et al. (2020). Transfer learning and fine tuning in mammogram BI-RADS classification. In 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS). IEEE.

Hong, Y., He, J., Deng, D., Liu, Q., Zu, X., & Shen, Y. (2025). Targeting kinases that regulate programmed cell death: A new therapeutic strategy for breast cancer. Journal of Translational Medicine, 23(1), 439.

Hussain, D., Al-Masni, M. A., Aslam, M., Sadeghi-Niaraki, A., Hussain, J., Gu, Y. H., & Naqvi, R. A. (2024). Revolutionizing tumor detection and classification in multimodality imaging based on deep learning approaches: Methods, applications and limitations. Journal of X-Ray Science and Technology, 32(4), 857–911.

Kaur, P., & Mahajan, P. (2025). Detection of brain tumors using a transfer learning-based optimized ResNet152 model in MR images. Computers in Biology and Medicine, 188, 109790.

Laishram, R., & Rabidas, R. (2024). Binary tunicate swarm algorithm based novel feature selection framework for mammographic mass classification. Measurement, 235, 114928.

Lee, R. S., Gimenez, F., Hoogi, A., Miyake, K. K., Gorovoy, M., & Rubin, D. L. (2017). Data descriptor: A curated mammography data set for use in computer-aided detection and diagnosis research. Scientific Data, 4(1), 1–9.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., et al. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60–88.

Manjunatha, A., & Mahendra, G. (2024). TransNet: A hybrid deep learning architecture combining CNNs and transformers for enhanced medical image segmentation. In 2024 International Conference on Computing and Intelligent Reality Technologies (ICCIRT) (pp. 221–225). IEEE.

Mienye, I. D., Swart, T. G., Obaido, G., Jordan, M., & Ilono, P. (2025). Deep convolutional neural networks in medical image analysis: A review. Information, 16(3), 195.

Meenalochini, G., & Ramkumar, S. (2024). A deep learning based breast cancer classification system using mammograms. Journal of Electrical Engineering & Technology, 19(4), 2637–2650.

Ojo, O. A., Grant, S., Nwafor-Ezeh, P. I., Maduakolam-Aniobi, T. C., Akinborode, T. I., Ezenabor, E. H., & Ojo, A. B. (2025). Ferroptosis as the new approach to cancer therapy. Cancer Treatment and Research Communications, 100913.

Prusty, A., Singh, R., & Verma, K. (2022). Optimization of VGG16 for breast cancer classification on the MIAS dataset. Computational Imaging and Vision, 34(3), 101–115.

Potsangbam, J., & Devi, S. S. (2024). Classification of breast cancer histopathological images using transfer learning with DenseNet121. Procedia Computer Science, 235, 1990–1997.

Salehi, A. W., Khan, S., Gupta, G., Alabduallah, B. I., Almjally, A., Alsolai, H., et al. (2023). A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope. Sustainability, 15(7), 5930.

Santos, J. C., Santos, M. S., & Abreu, P. H. (2024). Enhancing mammography: A comprehensive review of computer methods for improving image quality. Progress in Biomedical Engineering.

Seneng, I. K., Ayu, P. D. W., & Huizen, R. R. (2025). Comparative analysis of augmentation and filtering methods in VGG19 and DenseNet121 for breast cancer classification. Jurnal Teknik Informatika (Jutif), 6(3), 1131–1146.

Sharma, P., Gupta, R., & Mishra, S. (2022). DenseNet for malignant breast anomaly classification: Insights and challenges. Histopathology Insights, 28(1), 23–39.

Takahashi, S., Sakaguchi, Y., Kouno, N., Takasawa, K., Ishizu, K., Akagi, Y., et al. (2024). Comparison of vision transformers and convolutional neural networks in medical image analysis: A systematic review. Journal of Medical Systems, 48(1), 84.

Taifi, K., Sabbar, Y., Ardah, H., & Abdel-Aty, A. H. (2025). Enhanced DL-based breast cancer diagnosis and classification using modified DenseNet-121, DenseNet-201, and MobileNetV2: Optimized architectures and refined activation functions. IEEE Access.

Wang, J., Li, B., Luo, M., Huang, J., Zhang, K., Zheng, S., et al. (2024). Progression from ductal carcinoma in situ to invasive breast cancer: Molecular features and clinical significance. Signal Transduction and Targeted Therapy, 9(1), 83.

Xiong, X., Zheng, L. W., Ding, Y., Chen, Y. F., Cai, Y. W., Wang, L. P., et al. (2025). Breast cancer: Pathogenesis and treatments. Signal Transduction and Targeted Therapy, 10(1), 49.

Yiallourou, A. I. (2023). Hereditary breast cancer syndromes. In Breast Cancer Management for Surgeons: An Examination Guide (p. 79).

Zahra, N., Ahmed, H., & Malik, S. (2023). Deep learning-based multi-network feature fusion for breast cancer detection. Artificial Intelligence in Medicine, 120(4), 1–12.