Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 6, 2026

ADVANCEMENTS IN BREAST CANCER DETECTION: AN ANALYSIS OF PRE-TRAINED CNN MODELS AND TRANSFORMER-BASED TRANSFER LEARNING ON MAMMOGRAM SLIDES

[1]Azqa Fatima, [2]Ghazanfar Ali, [*3]Salahuddin, [4]Hafsa Hussain, [5]Nasir Hussain
[1]Department of Computer Science & IT, University of Southern Punjab Multan, Pakistan.
[2]Department of Computer Science, University of South Asia, Lahore, Pakistan.
[*3]Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan.
[4]Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan.
[5]Department of Computer Science & IT, University of Southern Punjab Multan, Pakistan.
[*3]msalahuddin8612@gmail.com

DOI:-https://doi.org/10.5281/zenodo.21258060

Keywords
Breast cancer detection; Malignant tumors; INbreast mammogram dataset; Transfer learning; Hierarchical Attention Networks (HAN); Transformer encoder layer

Article History
Received: 25 May, 2026
Accepted: 29 June, 2026
Published: 30 June, 2026

Corresponding Author: * Salahuddin

Abstract
Breast cancer remains a formidable disease, claiming millions of lives worldwide annually. Timely and precise detection of malignant tumors is paramount to enhancing patient prognosis. In our investigation, we distinguish between benign and malignant tumors and evaluate the efficacy of fourteen pre-trained convolutional neural network (CNN) models using the INbreast cancer dataset. Leveraging transfer learning methodologies to discern malignant tumors across 410 mammograms. In a pioneering, we conduct a novel analysis of attention weight scores generated by Hierarchical Attention Networks (VIT-L16) within a deep dense transfer network to discern the most influential factors in predicting malignant tumors. Our methodological approach encompasses a range of evaluation metrics for all pre-trained models, laying the groundwork for the development of an automated system proficient in identifying various breast lesions. The integration of transformers with transfer learning presents several benefits. Firstly, transformers excel at capturing long-range dependencies within data, enabling them to effectively analyze intricate patterns and relationships present in medical images, such as those in mammograms. Secondly, by leveraging pre-trained transformer models, we can expedite the training process and mitigate the need for vast amounts of labeled data, thus facilitating the development of accurate diagnostic systems even in scenarios where labeled medical datasets are scarce. Ultimately, the fusion of transformers with transfer learning equips us with powerful tools to enhance the early and accurate detection of breast cancer, with profound implications for the realms of medical imaging and computer-aided diagnosis (CAD). Our experimental findings reveal ResNet50 as the top-performing model, achieving an accuracy of 73.09%.

1. Introduction

Cancer is a devastating disease that poses a significant threat to human life. Breast cancer is one of the most common types of cancer, affecting millions of women worldwide. [1] According to the World Health Organization (WHO), breast cancer is the leading cause of cancer death among women worldwide, with an estimated 2.3 million new cases and 685,000 deaths in 2020 alone [2]. Several risk factors are associated with breast cancer, including age, gender, family history, genetic mutations, and lifestyle factors such as alcohol consumption and obesity. Early detection of breast cancer is crucial in reducing the mortality rate associated with this disease, enabling timely and effective treatment [3–5]. In addition, the five-year survival rate for localized breast cancer that has not spread to other parts of the body is as high as 99%, highlighting the importance of early detection [6–8].

Unfortunately, the current method of diagnosing breast cancer relies heavily on mammography images, an X-ray-based imaging technique that produces two-dimensional images of the breast tissue. Mammography is effective at detecting breast cancer at an early stage, but it has several limitations [9,10,10]. For example, mammography images can be affected by impurities, noise, and low resolution can affect mammography images, making it difficult to diagnose breast cancer accurately. Furthermore, the subjective nature of medical practitioners can lead to missed diagnoses or false positives, resulting in unnecessary treatments and procedures.

To overcome these limitations, researchers have proposed using CAD systems that can assist clinicians and radiologists in the early-stage detection of breast cancer using mammographic images with greater accuracy than traditional methods [11–13]. CAD systems use image processing algorithms and machine learning techniques to analyze mammography images and detect and identify abnormalities or suspicious regions that may be indicative of breast cancer. As a result, CAD systems can help reduce the number of false negatives and false positives, improving patient outcomes.

The accuracy of CAD systems depends on the quality of mammography images and the selection of appropriate image features. Several studies have shown that the performance of CAD systems can be improved by using deep learning (DL) techniques, which can automatically extract features from mammography images and classify them as benign or malignant [14,15]. In recent years, DL models have gained popularity among medical practitioners due to their automatic feature selection properties. However, DL models require large amounts of training data to achieve high accuracy, and training them can be computationally expensive [16].

To address these challenges, researchers have proposed using transfer learning (TL), which has demonstrated significant improvements over convolutional neural network (CNN) models. TL involves using a pre-trained model, such as the AlexNet model, which has been trained on a large amount of image data, to classify new images accurately.

Furthermore, it allows researchers to reduce the computational cost of training a DL model from scratch and enables knowledge transfer from one domain to another. Many existing approaches use the TL preset hyper-parameters, which have been chosen as the industry standard. However, recent studies have highlighted the need for customized hyperparameter tuning to achieve better results.

The goal of this research paper is to investigate the effectiveness of DL-based CAD systems for the early detection of breast cancer using mammography images. In particular, we aim to evaluate the performance of TL-based DL models for breast cancer detection and compare them to traditional CAD systems and human radiologists. We will also explore the impact of hyperparameter tuning on the performance of DL models and investigate the feasibility of deploying these models in real-world clinical settings.

- We assessed fourteen pre-trained models, each enhanced with two additional layers tailored for categorizing cancer into the BC category. These layers were interconnected using a global average pooling layer and fully connected through binary cross-entropy and sigmoid functions.
- Our study pioneers the analysis of attention weight scores generated by Hierarchical Attention Networks (VIT-L16) within a deep dense transfer network. Our aim is to identify crucial factors influencing prediction outcomes and model success, offering valuable insights for both researchers and practitioners.

- We validated the Transformer's performance in computer vision, especially in attention mammogram analysis (VIT-L16 Transformer), resulting in improved accuracy for computer vision tasks. The benefit of using Transformers with transfer learning lies in their ability to leverage pre-trained representations of data, enabling faster convergence and requiring less labeled data for training.
- Finally, through a performance-driven analysis of fourteen deep transfer learning models, we have developed highly accurate automatic BC prediction models, promising advancements in breast cancer diagnosis and treatment strategies.

2. Methodology

In this section, we present a comprehensive exposition of the BC INbreast mammograms dataset, outlining each preprocessing step and a range of deep transfer learning network tasks. In addition, we have meticulously documented the earlier deep learning models, which were subsequently reconstructed and subjected to a comparative analysis based on standardized model hyperparameters, such as the number of epochs, learning rate, batch size, and input image size (224 x 224). Figure 1 depicts the architecture of the suggested methodology.

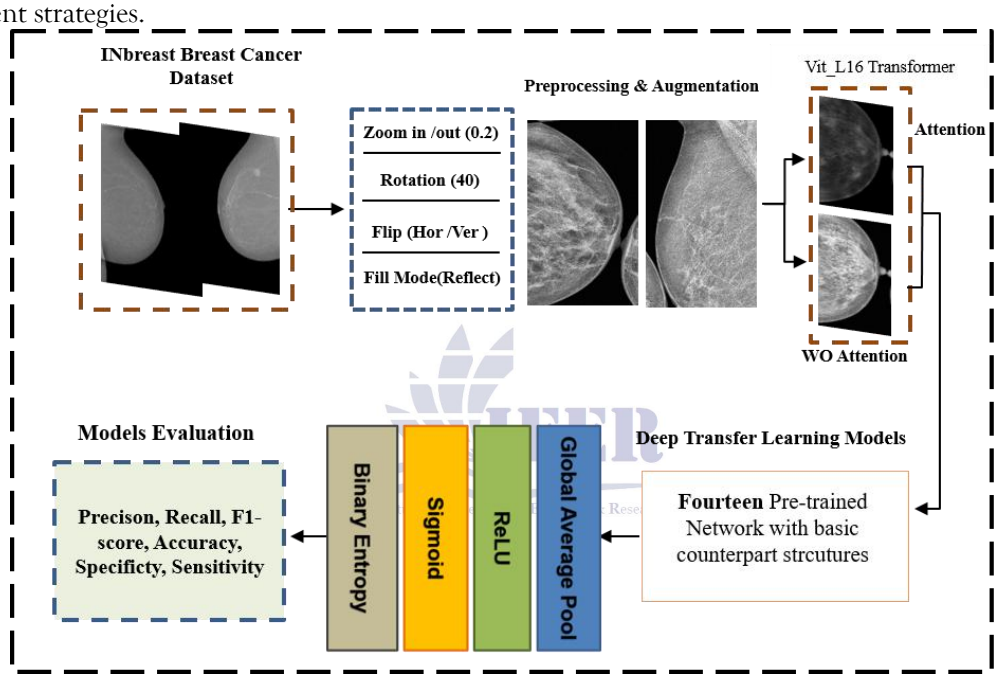

Figure 1. Propose Methodology Diagram

2.1. Dataset Details

In this research, we utilized the INbreast mammography slides as a foundational dataset, comprising 410 BC mammogram slides classified into two principal categories: benign and malignant. Table 1 provides a detailed breakdown of the data set, while Figure 2 illustrates the categories of sample slides for each classification. The INbreast dataset comprises mammography slides from the medical history of 115 patients, with a 14-bit pixel dimension and two views, MLO (Medio-Lateral-Oblique) and CC (Crania Caudal).

Table 1: *INbreast Dataset Details*

Properties	Details
root	Portugal
Lesion Category	All
Total Cases	115
Views	CC- (Left, Right) and MLO- (Left, Right)
Total Mammograms	410
Resolution	14 bits (Pixel)
Ratio (Malignant: Benign)	0.28:0.72
Annotation	finding label level through annotation

Breast Density Details	Yes

2.2. Preprocessing

The stratified [17] balancing approach and the UnSharp mask with smooth mask image enhancement technique are employed in our proposed methodology to normalize mammography slides. As a result, the model has minimized unbiased learning towards minority class samples in the INbreast mammograms dataset while also guaranteeing good generalization. This study has employed three typical transfer learning steps. First, normalize the distorted mammograms, then balance the training dataset using a stratified [17] balancing strategy, and finally, lower the likelihood of model overfitting in transfer learning. The slides have been enhanced with the slide image (224x224) dimension and sent to pre-trained models [18–22]. The ratio of train 80: valid 10: test 10 was used to partition the BC Cancer process INbreast dataset.

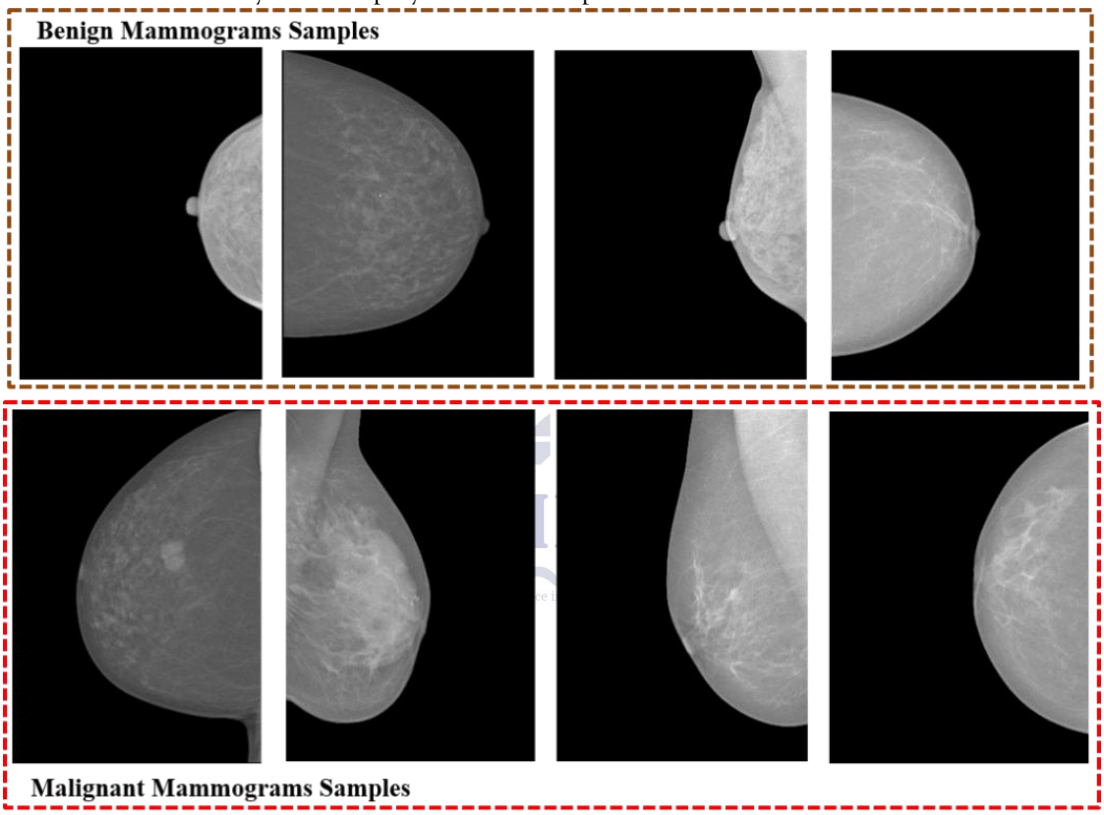

Figure 2. Sample of Benign Mammograms and Malignant Mammograms

2.3. Data Augmentation

During the training phase, we used various sets of random augmentation flips. We divided the INbreast dataset into smaller image sizes (224 x 224) W x H, slide rotation (400), and flip (Vertical, Horizontal) for the input image size to the network. Figure 3 displays enhanced augmented slides and Table 2 is presenting the total number of samples before and after augmentation.

Table 2: *Before and After Augmentation INbreast Dataset Details*

	Original Samples		Augmented Samples	
Dataset	Benign	Malignant	Benign	Malignant
INbreast	302	108	1493	1488
	Total Samples 410		Total Samples 2981	

2.4. Attention Mammogram (Vit-L16 Transformer)

The authors of [23] described a transformation architecture that combined the advantages of natural language processing and computer vision. The three layers that make up this design are the layer of self-attention layer, the layer encoder-decoder, and the position-wise levels (feed-forward) along the Relu activation function. Each pixel channel in the encoder block input data in a contextual presentation. Further, the decoder block 124 and the auto-regressive channel for each pixel are congruent. The key-value pairs with 125 the query, using a self-attention layer. Transformers are the only ones that employ self- 126 attention for sequence alignment as opposed to RNNs and CNNs. Building architecture 127 based on transformers is still being developed. In this study, we validate the attention base 128 mammogram with deep transfer learning models.

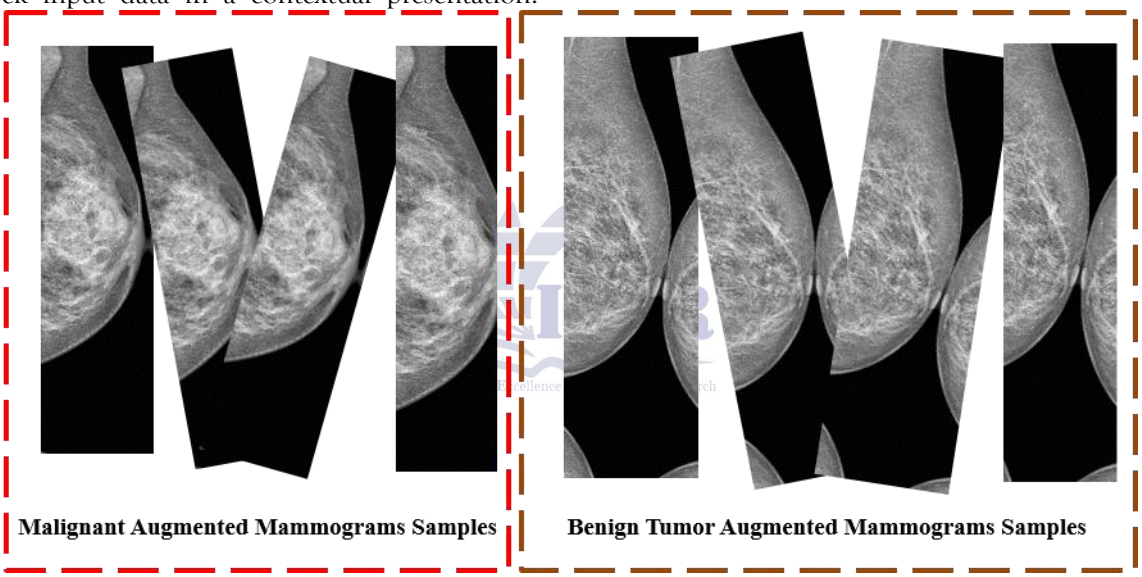

Figure 3. Samples of Augmented Mammograms

2.5. AlexNet

In [24], the AlexNet architecture has been divided into two layers, where five layers are related to the features layer (convolution layers), and three layers belong to the fully connected layer. In each layer, the ReLU activation function with filters of convolution has passed. The size of input slides is 224 x 224 with RGB channel due to fully connected layers. The 4096 neurons contain an FCN layer with a cross-entropy layer. Where each maintains a different feature for a different image. To handle the model overfitting problem, the dropout layer is applied at regular intervals.

2.6. DenseNet121

The features of all previous levels cannot be combined in this [25] model. These feature maps have been connected and utilized as inputs on the other side. In addition to two convolution feature layers with kernel sizes of 1x1 as a layer of congestion and a 3x3 kernel layer for executing convolution operations, the model's dense blocks each have a different number of block layers. a layer of Convolution 1x1, a layer of average pool layer 2x2 feature reduction with strides of 2 consists of layers of transition. In conclusion, stride 2 in the fundamental convolution layer 7x7 64 filters. Six times of dense block 1 are repeated together with stride 2 in the basic pool layer, layer 3x3 feature reductions layer and layer 2 convolution layers. The transitional layer consists of Convolution (1), and Average Pool

(1). Finally, all feature maps have been processed to carry out classification tasks using the global feature layer.

2.7. DenseNet201

There are four distinct densities of the dense layer in the model [26] The final dense block with classification layer, feature reduction 3x3 max pool layer with stride 2, and accept all the prior feature layers' maps of the DensNet201 network. There is a bottleneck layer for each convolution layer action. This layer caused the 1x1 convolutional layer to lower the number of input channels, while the 3x3 convolutional layer maintained the convolution process with a t input shape transform while maintaining the same input size and fewer input channels.

2.8. GoogleNet

The GoogleNet [27] pre-trained model has 9 inception modules along 22 convolution layers in its architecture. Three different sizes of the kernel (1 x 1, 3 x 3, 5 x 5) have used in the inception module, for pooling layer 3 x 3 filter size, and the size of the convolution filter is 1 x 1. Due to this small GoogleNet structure, reduce the computation cost as well as handle the depth of the feature map. The size of the input slide is 224 x 224 with the RGB color channel image. The stochastic gradient descent algorithm has been used to extract the most appealing feature from slides during the training phase through the structure of the GoogleNet network.

2.9. InceptionResNetV2

In this [28] deep transfer learning model architecture residual blocks combine with inception blocks, and a total of 164 deep hierarchical networks consist in it. The input size of slides should be 229 x 229 pass-through in this network. The stem is the main block of this network proceeded by Inception with A-resent, A-Reduction Inception B-resent, B-Reduction Inception, and C-resent modules. While the sequence of the convolution layer is 3x3, the layer of max-pool is 3x3, and filters combine to further process by the inception of the A-resent block via the main stem block. The input image in each inception block is through by activation function Relu, and the block of reduction retains the filter size of convolution 1x1, 3x3 with a layer of pooling. The batch normalization layer has been added to the top of the layers in the Inception ResNet block.

2.10. InceptionV3

To enhance the effectiveness of modules-based deep learning [28] introduced in the family of inception, many changes have been suggested e.g., Smoothing of Label, factorizing of 7x7 convolution, and label propagation combined with an additional classifier reduces the network size. The first convolution block features a reducing layer 3x3 with a stride of 2 applied and has an input size of 229 x 229, a kernel size of 3x3, and a stride of 2 after two. Three modules Inception (3) module 35x35, Inception (5) module 17x17, and Inception Module (2) 8x8 are configured before the categorization layer.

2.11. MobileNet

In [29] author proposed a simplified approach with separable depth-wise convolution layers for the lightweight application. In comparison to previous CNN models, the MobileNet model is capable of performing classification tasks in mobile and visual applications rapidly. Every input channel in this model has a single convolution layer, and each output channel uses a combination of the point convolution filter and a 1x1 convolution block before producing a depth-wise linear convolution.

2.12. NASNetMobile

Basic neural search net building blocks were used to create this [30] model, and reinforcement learning was then used to improve it. A sliding shape of 224 by 224 has been inserted into this model. The ability to regenerate these operations with a time number depends on the different convolution layers, pool procedures for feature reduction, and network depths. Twelve cells with accumulates of multiply exit this model. Two distinct cell structures are visible in the cell. The term "reduced cell" refers to a cell that reduces the height and breadth of the feature map by two factors while still producing a feature map with the same input dimension.

2.13. ResNet50

This model has a stem as its source of input. A single output layer is produced from four separate input phases [31]. 50 layers with a set image size make up the model; (22 4x 224). The stem input, 64 channels other than 2 pitch, and pass-through maximum pooling layer 3x3 a layer of feature reduction aside from 2 pitch were sent to the convolution layer 7x7. After lowering the image width and image height by 4 times by stem, the 64 channel's size was raised. The independent

stage has been used as the initial step in block-down sampling. Stage 2 has additionally begun following various iterations of the remaining blocks. There are now two active pathways in a block. Whereas A-path includes a 1x1 convolutional layer other than 2 pitch, A-path contains three convolutional layers 1x1, 3x3, and 1x1 kernel sizes, shifting the input to the output of path-A. The down-sampling block is added by A-path output and B-path output last.

2.14. ResNet101

Due to the dense residual network connections in a block, this [32] model has been extensively used in various biomedical applications for classification purposes. Following the resnet101 backbone layer, two forwarding 256x3 layers of convolutions are maintained by a normalization layer batch with a global pool layer that reduces the number of features. After the reducing layer of feature same, three convolutions with 128 3x3 layers each were added before the normalizing layer. Performance categorization has been applied with two flattened layers. Before normalizing layers, two dense layers with filter sizes (256, 64) were employed. Finally, a label dense layer of multiclass with a vector of probabilities derived from numbers.

2.15. ResNet152

To allow the trained model to learn a much deeper model (152 layers), ResNet152 [28] is a CNN's network base model that regularizes the functioning of the caper connections. A 224 by 224 slide image was provided to the model as input. A 3x3 feature that reduces the maximum pool layer with a stride of 2 has come after the convolution 64 of 7x7 with a stride of 2. Four convolution blocks with various layer architectures have been employed in this deeper model to carry out binary class and multi-classification tasks.

2.16. VGG16 & VGG19

The only deep learning model with an architecture that allows for decent local memory consumption on the GPU is the VGG16 [33] model. There are 16 layers in this design, 13 of which are convolution layers, 5 of which are feature-reducing max pool layers, and 3 of which are fully connected layers. These additional layers bring the total to 21 layers. There are only 16 layers of weight employed. RGB pixels in the color image with 224x224 3 information layer. The subsequent convolution layers have been performed with the Relu activation function 3x3 kernel size. The 19 trained layers that make up VGG19 [25]combine with various convolution layers, fully linked layers, a layer of feature-reduction, and dropout layers.

2.17. Xception

The Xception [34] model differs from the original Inception model primarily in two ways. The first is the order of operations; in the original, the first spatial convolution performs with a depth-wise convolution layer of separable and a channel-wise and 1x1 convolution layer. The second factor is the presence or absence of nonlinearity; in the original, nonlinearity would follow the first operation. While Xception uses a depth-wise separable modified convolution layer.

3. Experiments Details and Results

The field of medical image analysis has witnessed significant advancements in recent years, particularly in the application of deep learning techniques. One crucial area of research within this domain is the detection of breast cancer, which heavily relies on the accurate analysis of mammography images. This study aims to evaluate the performance of fourteen deep convolutional neural networks (CNNs) [31, 32] with and without attention mechanisms for the binary classification of benign and malignant breast tumors using mammography images.

The evaluation was conducted on the publicly available INbreast mammogram slides dataset, comprising 410 mammography slides. Each slide has a resolution of 3328 x 4084 pixels and a bit depth of 16. The dataset consists of two main types of tumors: benign and malignant, making it a binary classification problem. To enhance the dataset's quality and quantity, pre-processing techniques such as image enhancement and augmentation were employed.

Fourteen different CNN architectures were tested, each varying in terms of the number of layers and parameters. In order to optimize the models' performance, adjustments were made to the last layer, including modifications to the global average pool layer, dropout layer, cross-entropy for identifying the breast cancer category, and sigmoid activation. The same experimental settings [33-35] were applied to both CNNs with and without attention mechanisms, including a batch size of 64, 30 epochs, and an initial learning rate of 0.001.

To evaluate the performance of the models, the dataset was divided into three sections: 80% for training, 10% for validation, and 10% for testing. The classification report, which includes metrics such as accuracy, sensitivity, recall, precision, and f1-score, was used to assess the models' performance. While AlexNet demonstrated promising results in the binary classification of benign and malignant breast tumors on mammography images, achieving notable performance, it should be noted that its accuracy was relatively lower compared to other models evaluated in this study. AlexNet, with its deep architecture comprising multiple convolutional and fully connected layers, showed potential in distinguishing between the two tumor types. Its ability to classify between the two tumor types was evident in its classification results (as shown in Figure 4), while a detailed classification report for the AlexNet model is provided in Table 3.

However, the model has certain limitations, such as its relatively larger number of parameters compared to newer models, which can make it significantly computationally expensive and require a larger amount of training data. Furthermore, the early architecture of AlexNet may struggle to capture more intricate and subtle features present in medical images. Despite these drawbacks, Alex Net's performance in breast tumor classification warrants further exploration and optimization, considering its potential in medical image analysis.

The classification report in Table 3 provides valuable insights for researchers aiming to refine and improve the accuracy and performance of the model. By addressing its limitations and capitalizing on its strengths, AlexNet has the potential to significantly contribute to improving breast tumor classification in clinical settings.

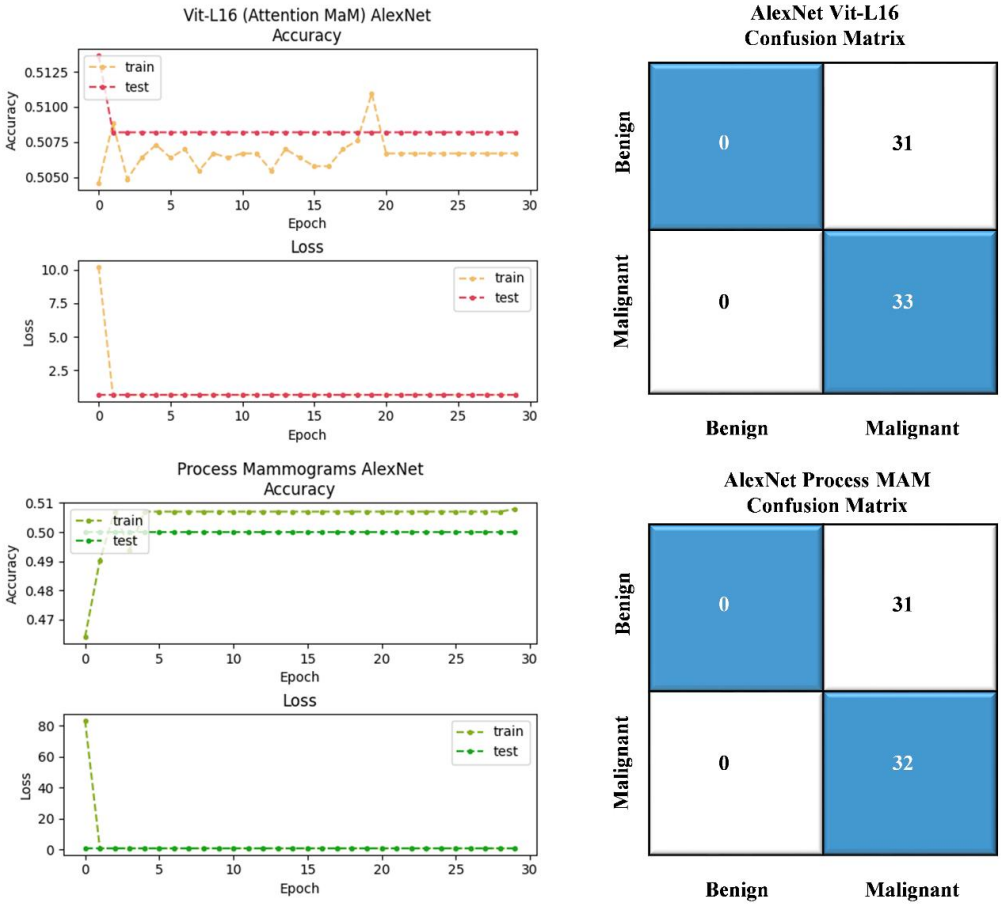

Figure 4. The training step ahead used pre-processed mammograms and attention VIT-L16 mammograms of AlexNet measurement during 30 Epochs Accuracy, Error Loss, Confusion Matrices

Table 3: *WO and With Attention Mammograms AlexNet Classification Report*

	Wo Attention		With Attention	
Metrics	Benign	Malignant	Benign	Malignant
Sensitivity	0.00	1.00	0.00	1.00
Specificity	0.00	0.70	0.00	0.69
Precision	0.00	0.51	0.00	0.51
Recall	0.00	1.00	0.00	1.00
F-Score	0.00	0.67	0.00	0.67
	Accuracy	50.79	Accuracy	50.68

Moving on to DenseNet121, this model presented a notable breakthrough in terms of accuracy and loss during training. Known for its dense connectivity pattern, DenseNet121 has the advantage of capturing complex patterns and features effectively. This architectural design promotes feature reuse and reduces the vanishing gradient problem. Furthermore, DenseNet121 demonstrated strong performance in classifying benign and malignant breast tumors, as evidenced by the accuracy results and confusion matrices. However, one potential drawback of DenseNet121 is its higher memory consumption due to its dense connectivity, which can limit its scalability and efficiency for larger datasets. Figure 5 has showing the performance breakthrough of DenseNet121 in terms of accuracy (Training, Validation), and Loss (Training, Validation) with 30 epochs.

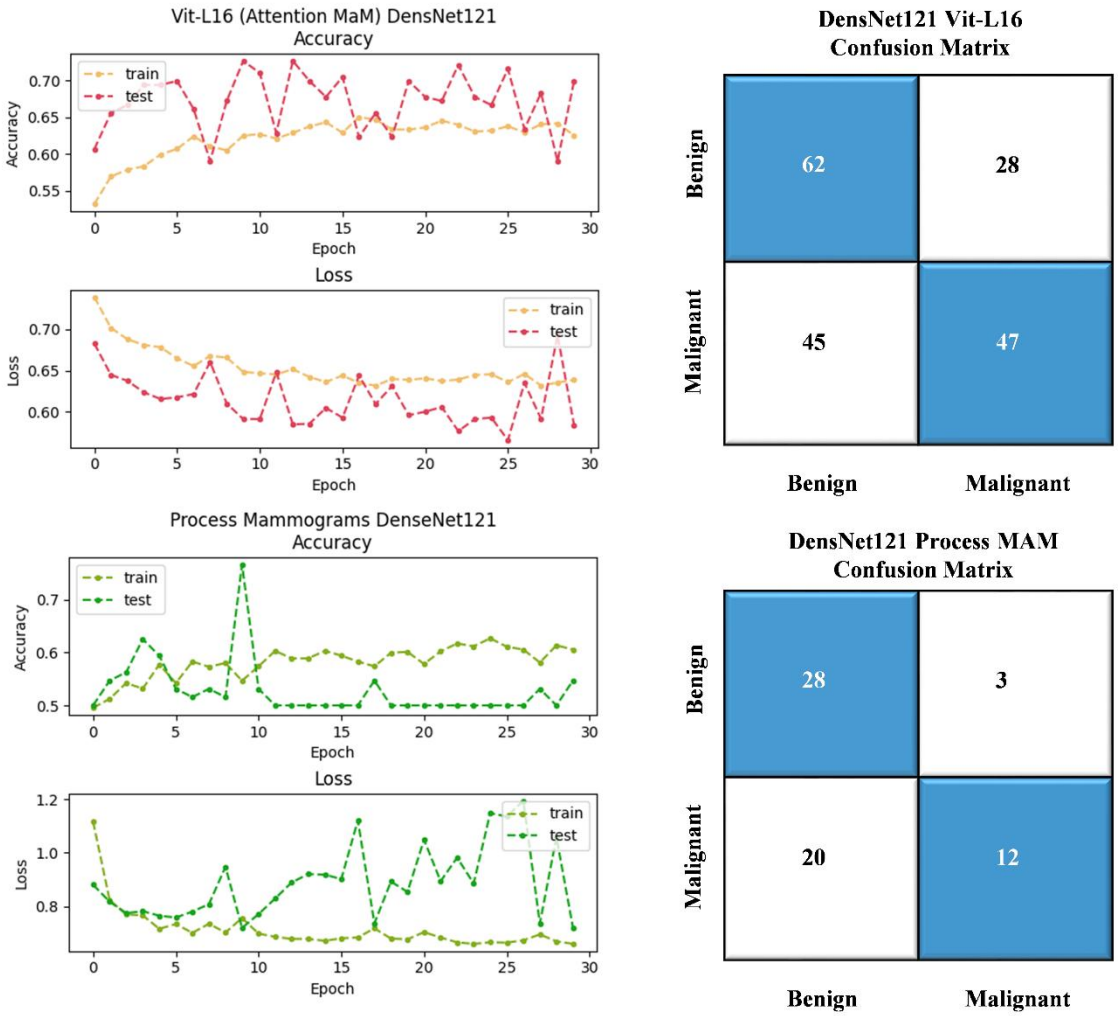

Figure 5. During 30 Epochs, the training Step forward using pre-process Mammograms and Attention VIT-L16 Mammograms of DenseNet121 measurement Accuracy, Error Loss, Confusion Matrices

The accuracy of each actual and predicted class is also represented in Figure 5 with confusion matrices. Table 4 shows the performance of the model with classification reports.

Table 4. *WO and With Attention Mammograms DenseNet121 Classification Report*

	WO Attention		With Attention	
Metrics	**Benign**	**Malignant**	**Benign**	**Malignant**
Sensitivity	0.90	0.38	0.69	0.51
Specificity	0.60	0.62	0.58	0.59
Precision	0.58	0.80	0.58	0.63
Recall	0.90	0.38	0.69	0.51
F-Score	0.71	0.51	0.63	0.56
	Accuracy	**63.49**	Accuracy	59.89

The next model evaluated, DenseNet201, showcased promising results in terms of accuracy and loss. It utilized both pre-processed mammograms and attention VIT-L16 mammograms, leveraging the power of attention mechanisms to enhance its performance. DenseNet201 demonstrated competitive accuracy in predicting the benign and malignant classes, as depicted in Figure 6.

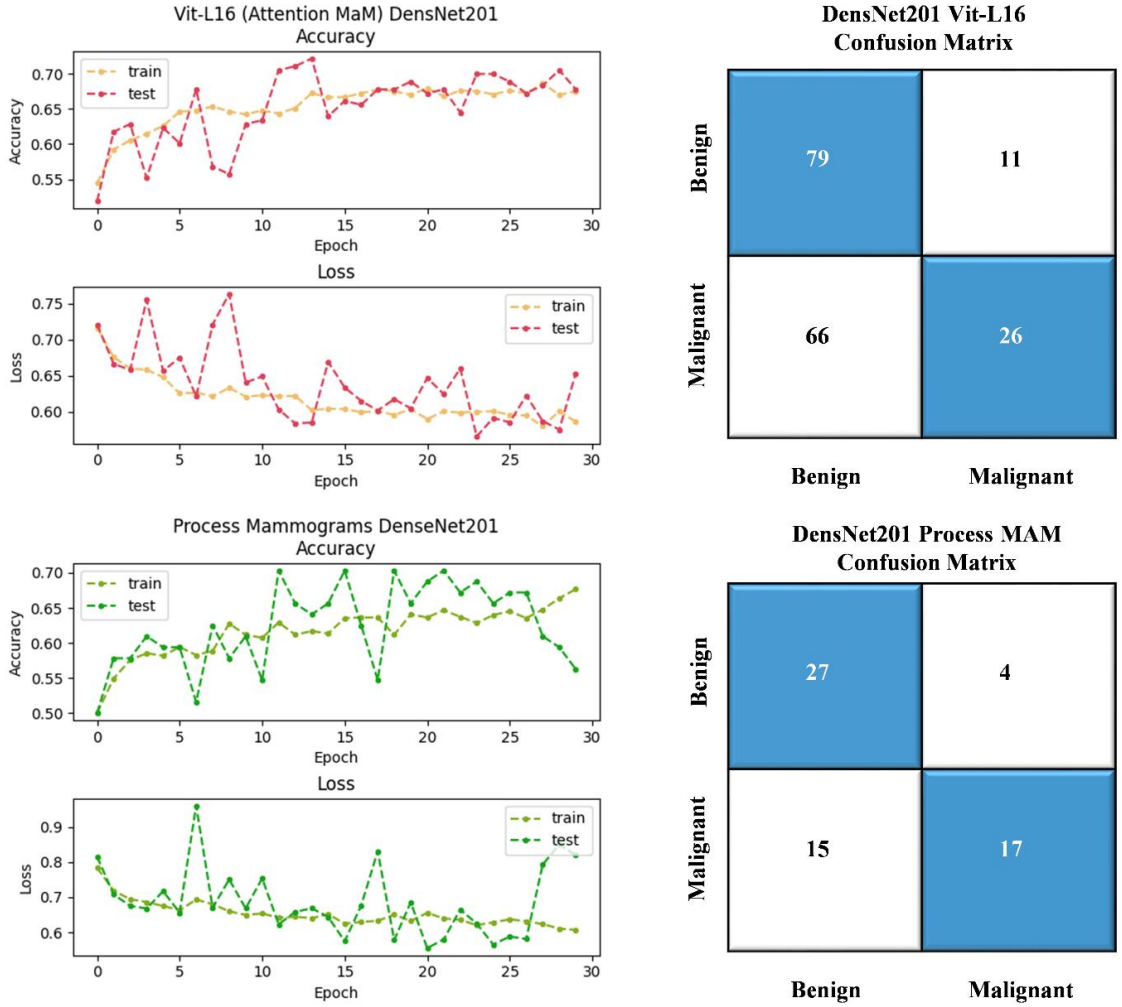

Figure 6. The training step ahead used pre-processed mammograms and attention VIT-L16 mammograms of DensNet201 measurement during 30 Epochs Accuracy, Error Loss, Confusion Matrices

The model classification report, presented in Table 5 further validates its improved performance. Despite its strong performance, DenseNet201, similar to DenseNet121, may face challenges related to memory consumption and

computational complexity due to its dense connectivity pattern.

Table 5. *WO and With Attention Mammograms DenseNet201 Classification Report*

	WO Attention		With Attention	
Metrics	Benign	Malignant	Benign	Malignant
Sensitivity	0.87	0.53	0.88	0.28
Specificity	0.63	0.65	0.56	0.54
Precision	0.64	0.81	0.54	0.70
Recall	0.87	0.53	0.88	0.28
F-Score	0.74	0.64	0.67	0.40
	Accuracy	**69.84**	Accuracy	57.69

GoogleNet, also known as InceptionV1, exhibited noteworthy performance in terms of accuracy and loss during training. This model introduced the concept of inception modules, which employ multiple filter sizes to capture features at different scales. The performance breakthrough of GoogleNet is depicted in Figure 7, where the accuracy and loss values are presented. The confusion matrices in Figure 7 provide insights into the accuracy of each class, both actual and predicted. While GoogleNet demonstrated competitive accuracy, it can be computationally expensive due to its complex architecture with multiple branches and filters. Table 6 shows model performance with classification reports.

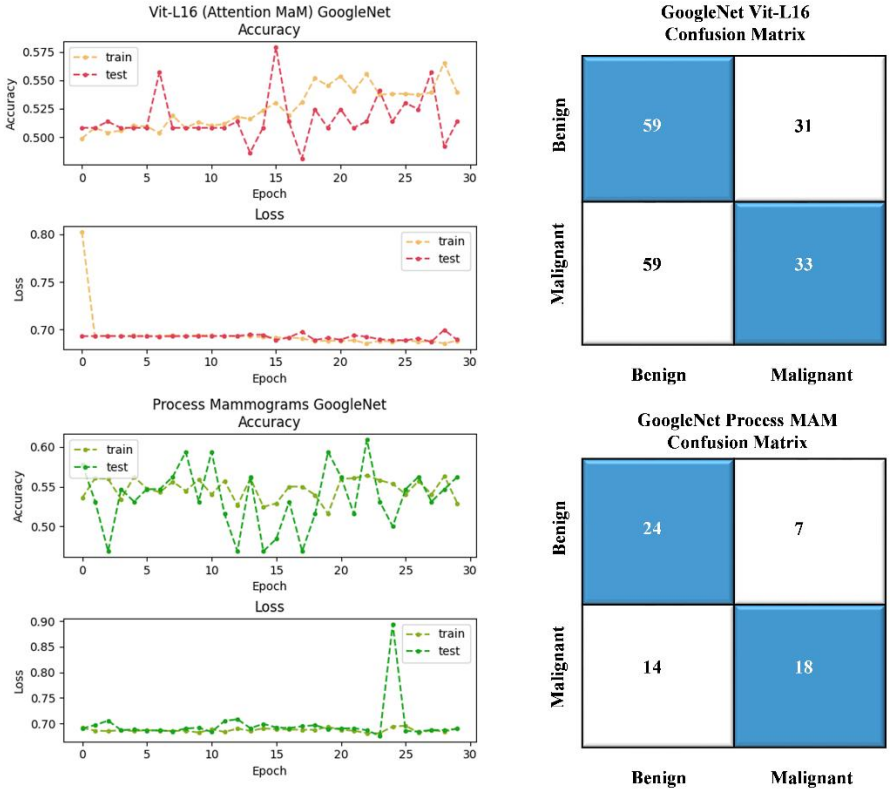

Figure 7. During 30 Epochs, the training Step forward using pre-process Mammograms and Attention VIT-L16 Mammograms of GoogleNet measurement Accuracy, Error Loss, Confusion Matrices

Table 6. *WO and With Attention Mammograms GoogleNet Classification Report*

	WO Attention		With Attention	
Metrics	Benign	Malignant	Benign	Malignant
Sensitivity	0.77	0.56	0.66	0.36
Specificity	0.65	0.64	0.50	0.49

Precision	0.63	0.72	0.50	0.52
Recall	0.77	0.56	0.66	0.36
F-Score	0.70	0.63	0.57	0.42
	Accuracy	66.66	Accuracy	50.54

InceptionResNetV2, a combination of Inception and ResNet architectures, demonstrated improved accuracy and performance in classifying benign and malignant breast tumors. This model, as shown in Figure 8, achieved promising results in terms of accuracy. The classification report presented in Table 7 further emphasizes its better performance. One advantage of InceptionResNetV2 is its ability to capture both local and global features effectively, due to its residual connections and inception modules. However, the trade-off for its improved performance is an increased number of parameters, which can impact computational efficiency.

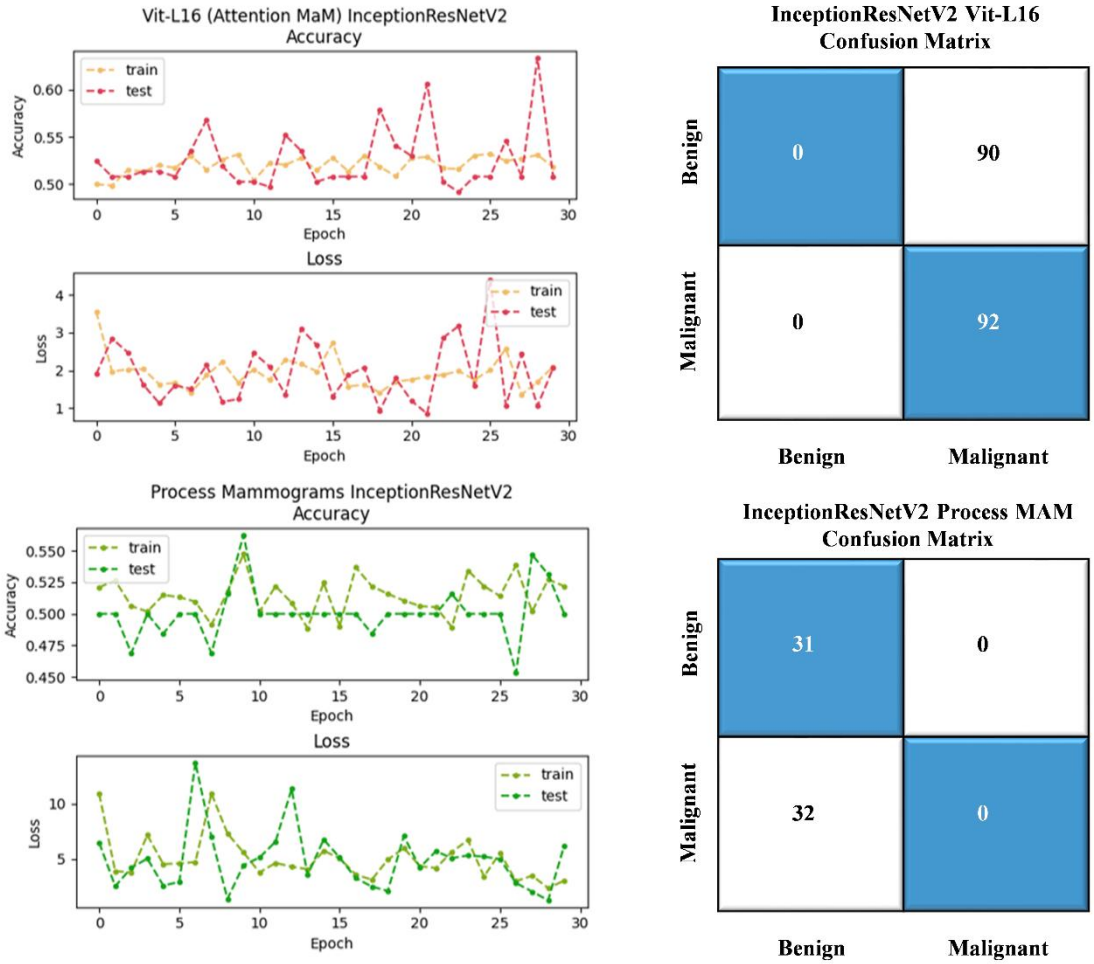

Figure 8. The training step ahead used pre-processed mammograms and attention VIT-L mammograms of InceptionResNetV2 measurement during 30 Epochs Accuracy, Error Loss, Confusion Matrices

Table 7: WO and With Attention Mammograms InceptionResNetV2 Classification Report

	WO Attention		With Attention	
Metrics	Benign	Malignant	Benign	Malignant
Sensitivity	1.00	0.00	0.00	1.00
Specificity	0.48	0.00	0.00	0.47
Precision	0.49	0.00	0.00	0.51
Recall	1.00	0.00	0.00	1.00

F-Score	0.66	0.00	0.00	0.67
	Accuracy	49.20	Accuracy	**50.54**

During 30 Epochs, the training Step forward using pre-process mammograms and attention VIT-L16 mammograms of Inceptionv3 measurement (Accuracy, Error Loss, Confusion Matrices) has been presented in Figure 9. The actual and predicted benign class or malignant class accuracy of Inceptionv3 has been presented in Figure 9. Furthermore, the classification report of the Inceptionv3 model is presented in Table 8.

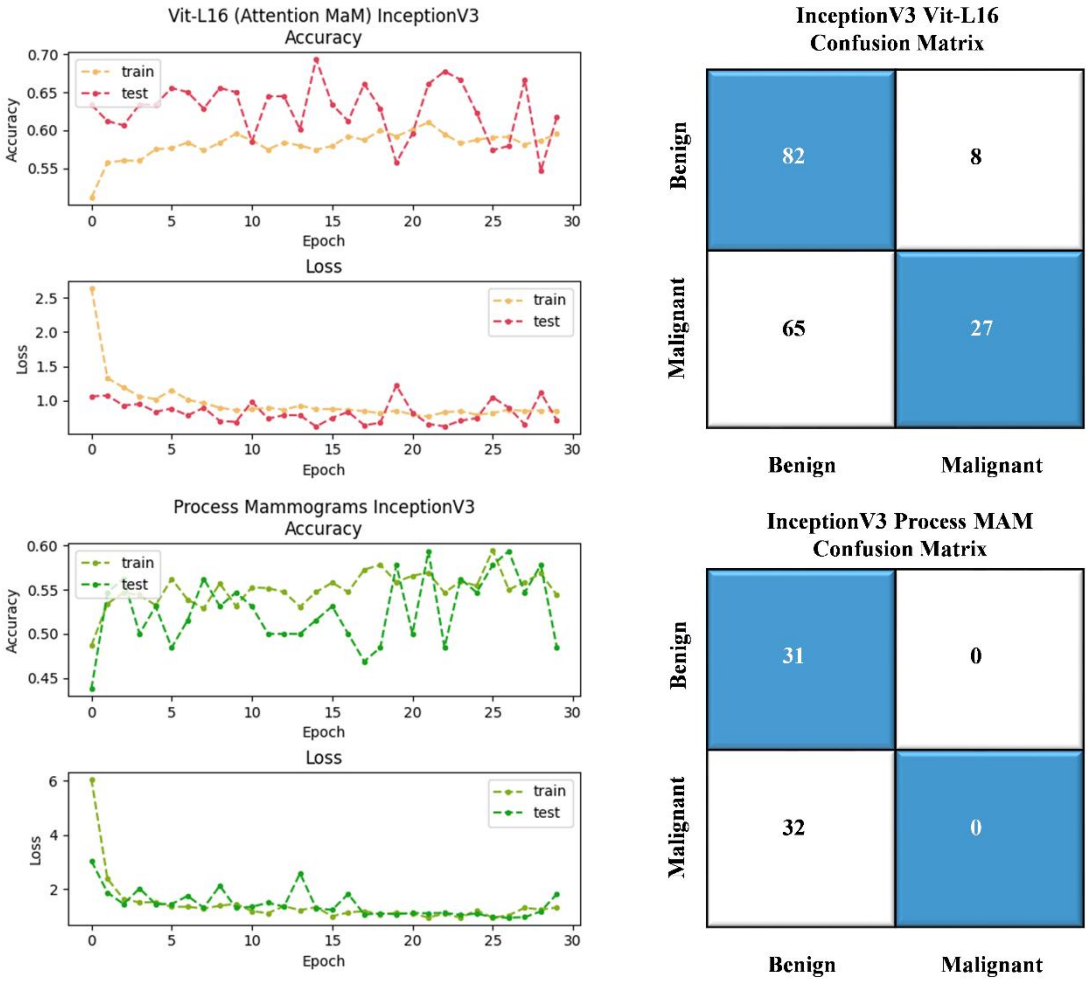

Figure 9. During 30 Epochs, the training Step forward using pre-process Mammograms and Attention VIT-L16 Mammograms of Inception V3 measurement Accuracy, Error Loss, Confusion Matrices

Table 8. *WO and With Attention Mammograms Inception V3 Classification Report*

	WO Attention		With Attention	
Metrics	**Benign**	**Malignant**	**Benign**	**Malignant**
Sensitivity	1.00	0.00	0.91	0.29
Specificity	0.45	0.00	0.56	0.57
Precision	0.49	0.00	0.56	0.77
Recall	1.00	0.00	0.91	0.29
F-Score	0.66	0.00	0.69	0.43
	Accuracy	49.20	Accuracy	**59.89**

MobileNet, a lightweight CNN model designed for mobile and resource-constrained devices, also showcased improved performance in classifying breast tumors. Figure 10 illustrates the accuracy of MobileNet in predicting the benign and malignant classes. With its depth wise separable

convolutional layers, MobileNet achieved competitive accuracy while maintaining a smaller model size and lower computational complexity. This makes it suitable for deployment on mobile platforms. However, MobileNet may struggle to capture more complex and fine-grained features due to its lightweight design.

Table 9 gives the classification report for the MobileNet model.

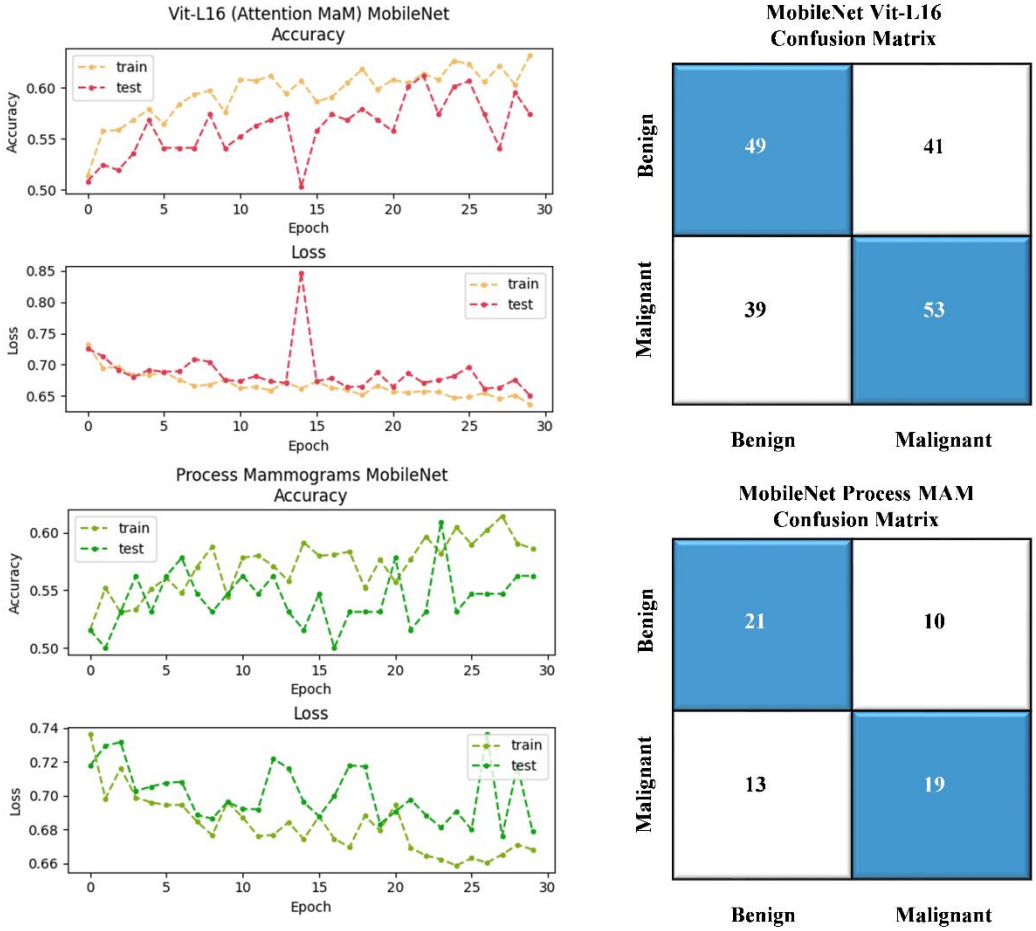

Figure 10. The training step ahead used pre-processed mammograms and attention VIT-L16 mammograms of MobileNet measurement during 30 Epochs Accuracy, Error Loss, Confusion Matrices

Table 9. *WO and With Attention Mammograms MobileNet Classification Report*

	WO Attention		With Attention	
Metrics	**Benign**	**Malignant**	**Benign**	**Malignant**
Sensitivity	0.68	0.59	0.54	0.58
Specificity	0.62	0.61	0.54	0.55
Precision	0.62	0.66	0.56	0.56
Recall	0.68	0.56	0.54	0.58
F-Score	0.65	0.62	0.55	0.57
	Accuracy	**63.49**	**Accuracy**	**56.04**

NASNetMobile, a neural architecture search-based model, exhibited a breakthrough in terms of accuracy and loss during training. Its performance, as depicted in Figure 11, demonstrates competitive accuracy in classifying benign and malignant breast tumors. NASNetMobile advantage lies in its ability to automatically discover a highly efficient and effective architecture through neural architecture search. However, this comes at the cost of increased computational complexity and resource requirements during the architecture search process. Table 10 shows model performance with classification reports.

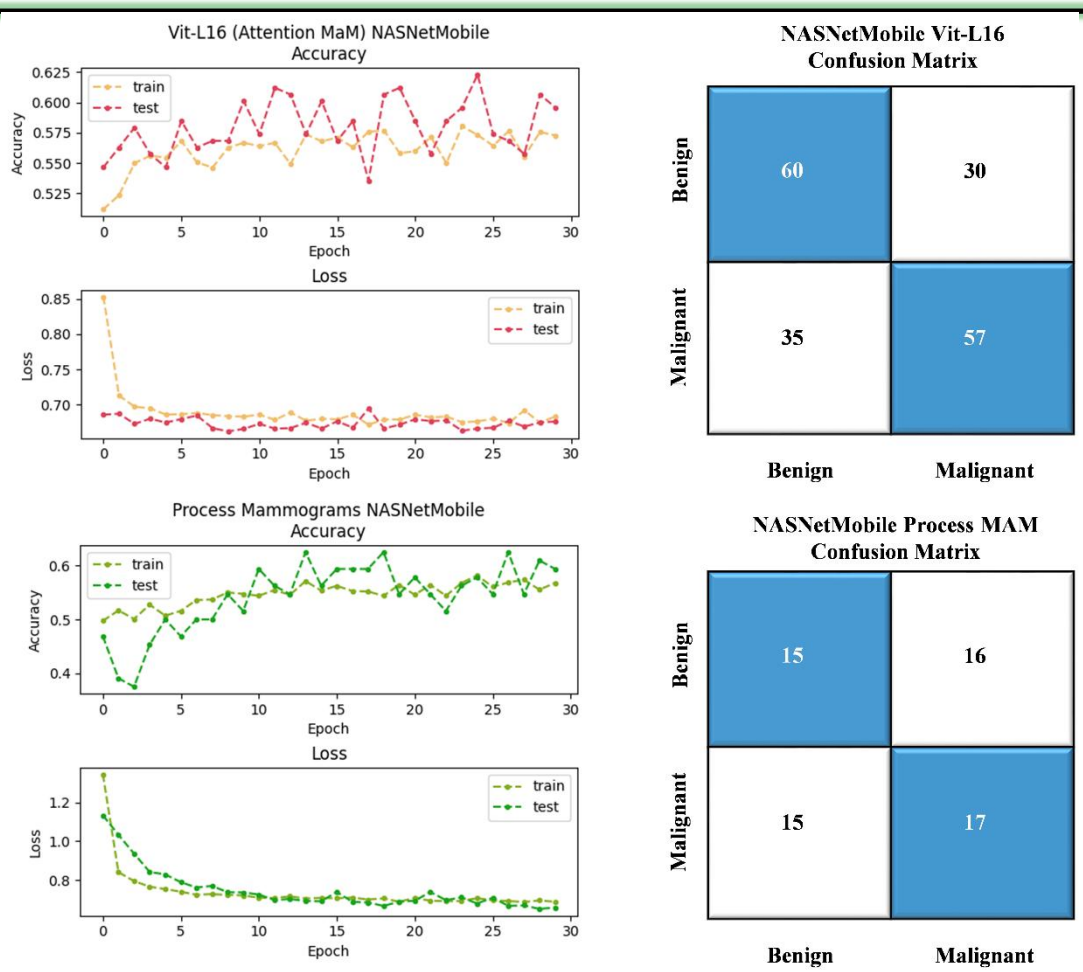

Figure 11. During 30 Epochs, the training Step forward using pre-process Mammograms and Attention VIT-L16 Mammograms of NASNetMobile measurement Accuracy, Error Loss, Confusion Matrices

Table 10. *WO and With Attention Mammograms NASNetMobile Classification Report*

	WO Attention		With Attention	
Metrics	**Benign**	**Malignant**	**Benign**	**Malignant**
Sensitivity	0.48	0.53	0.67	0.62
Specificity	0.47	0.49	0.63	0.62
Precision	0.50	0.52	0.63	0.66
Recall	0.48	0.53	0.67	0.62
F-Score	0.49	0.52	0.65	0.64
	Accuracy	63.49	Accuracy	**64.28**

ResNet50 and ResNet101, both belonging to the ResNet family, achieved impressive accuracy in classifying breast tumors. As shown in Figures 12 and 13, these models showcased their effectiveness in capturing important features and achieving high accuracy.

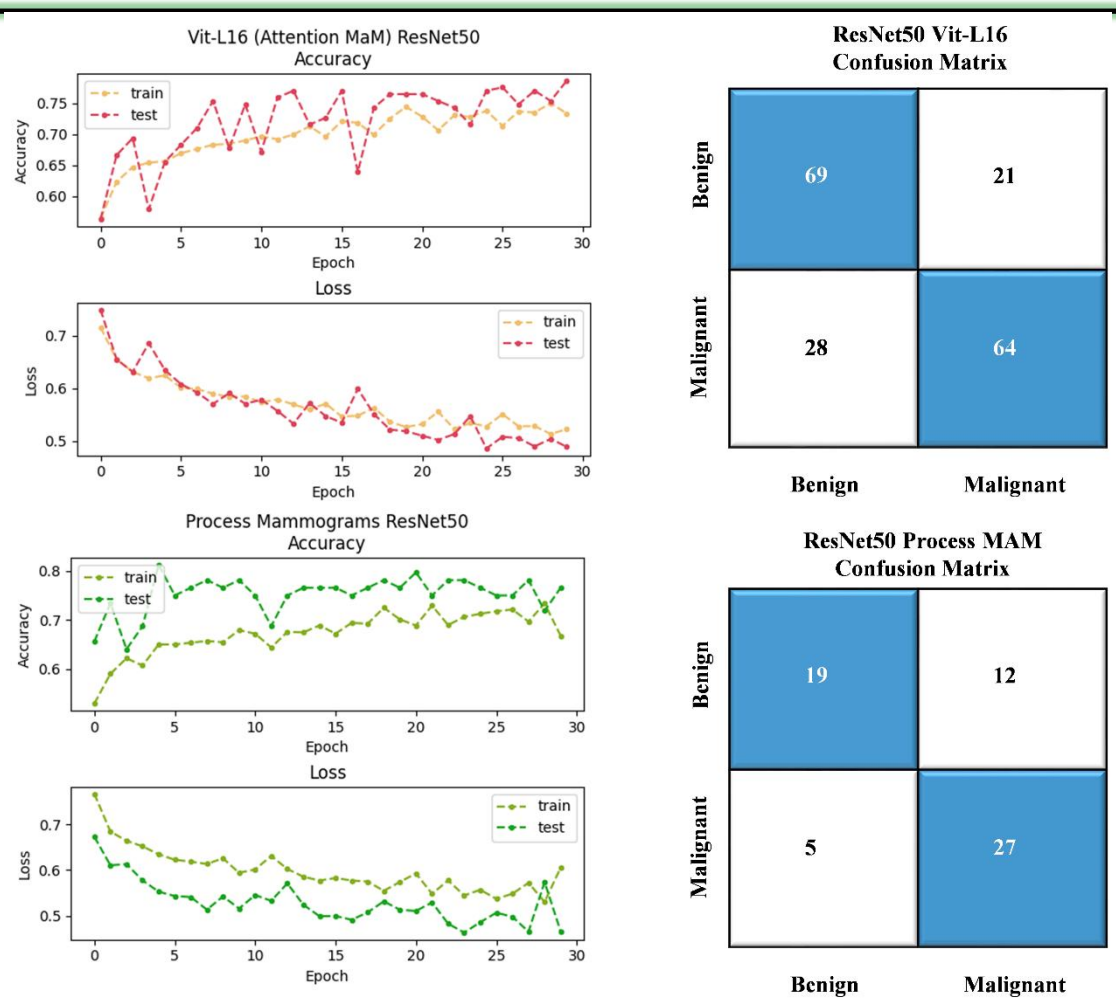

Figure 12. The training step ahead used pre-processed mammograms and attention VIT-L16 mammograms of ResNet50 measurement during 30 Epochs Accuracy, Error Loss, Confusion Matrices

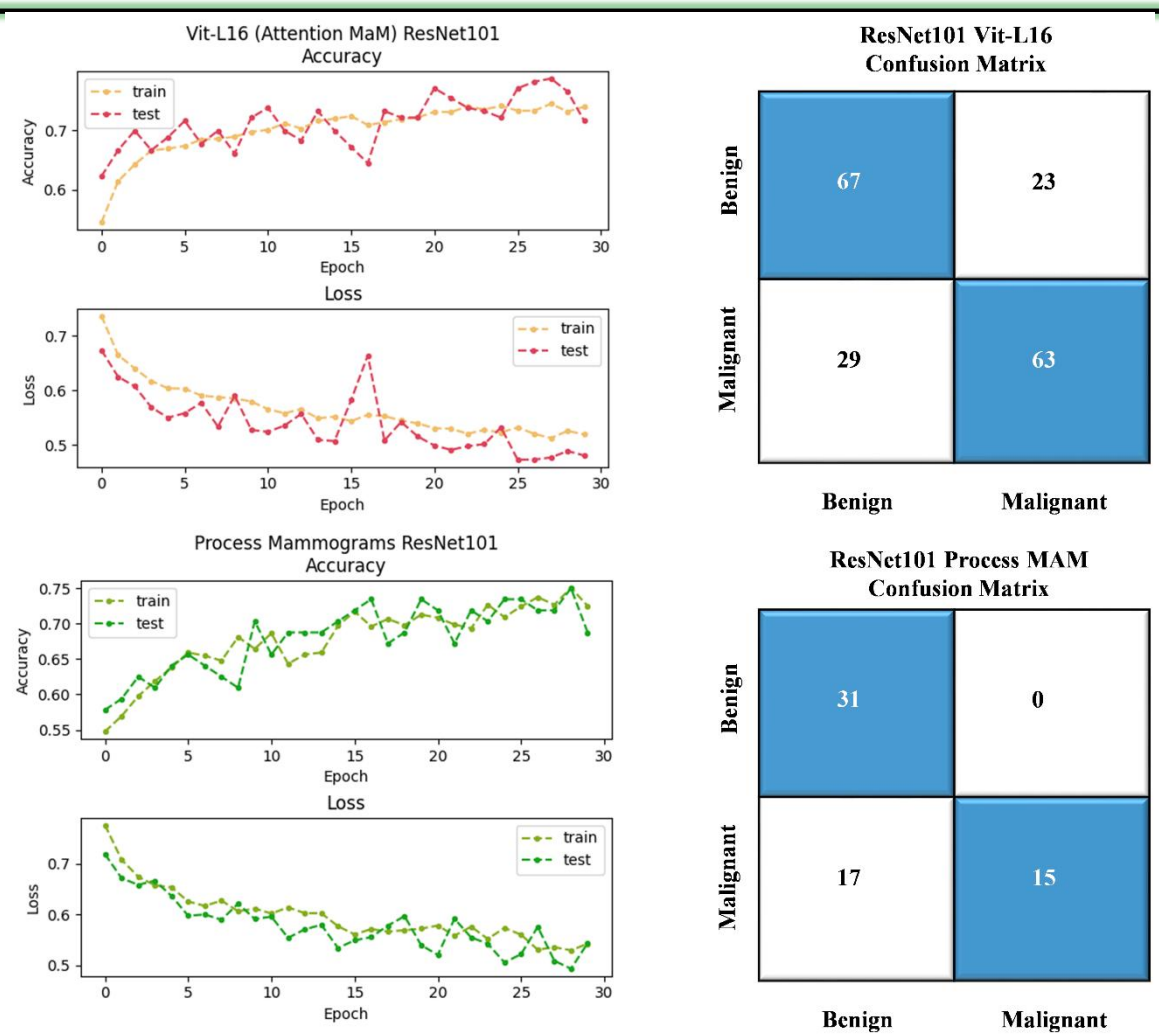

Figure 13. During 30 Epochs, the training Step forward using pre-process Mammograms and Attention VIT-L16 Mammograms of ResNet101 measurement Accuracy, Error Loss, Confusion Matrices

The classification reports in Tables 11 and 12 further validate their performance. ResNet models are renowned for their residual connections, which mitigate the vanishing gradient problem and facilitate the training of deeper networks. However, deeper ResNet models can be computationally demanding and may require substantial computational resources.

Table 11. *WO and With Attention Mammograms ResNet50 Classification Report*

	WO Attention		With Attention	
Metrics	Benign	Malignant	Benign	Malignant
Sensitivity	0.61	0.84	0.77	0.70
Specificity	0.60	0.65	0.65	0.61
Precision	0.79	0.69	0.71	0.75
Recall	0.61	0.84	0.77	0.70
F-Score	0.69	0.76	0.74	0.72
	Accuracy	73.01.49	Accuracy	**73.07**

Table 12. *WO and With Attention Mammograms ResNet101 Classification Report*

	WO Attention		With Attention	
Metrics	Benign	Malignant	Benign	Malignant
Sensitivity	1.00	0.47	0.74	0.68
Specificity	0.71	0.65	0.64	0.62
Precision	0.65	1.00	0.70	0.73
Recall	1.00	0.47	0.74	0.68
F-Score	0.78	0.64	0.72	0.71
	Accuracy	73.01.49	Accuracy	71.42

ResNet152, another member of the ResNet family, demonstrated a notable performance breakthrough in terms of accuracy and loss. As depicted in Figure 14, ResNet152 achieved high accuracy in classifying benign and malignant breast tumors. The model's classification report, presented in Table 13, confirms its impressive performance. Similar to other ResNet models, ResNet152 benefits from its residual connections but may have increased computational demands due to its depth.

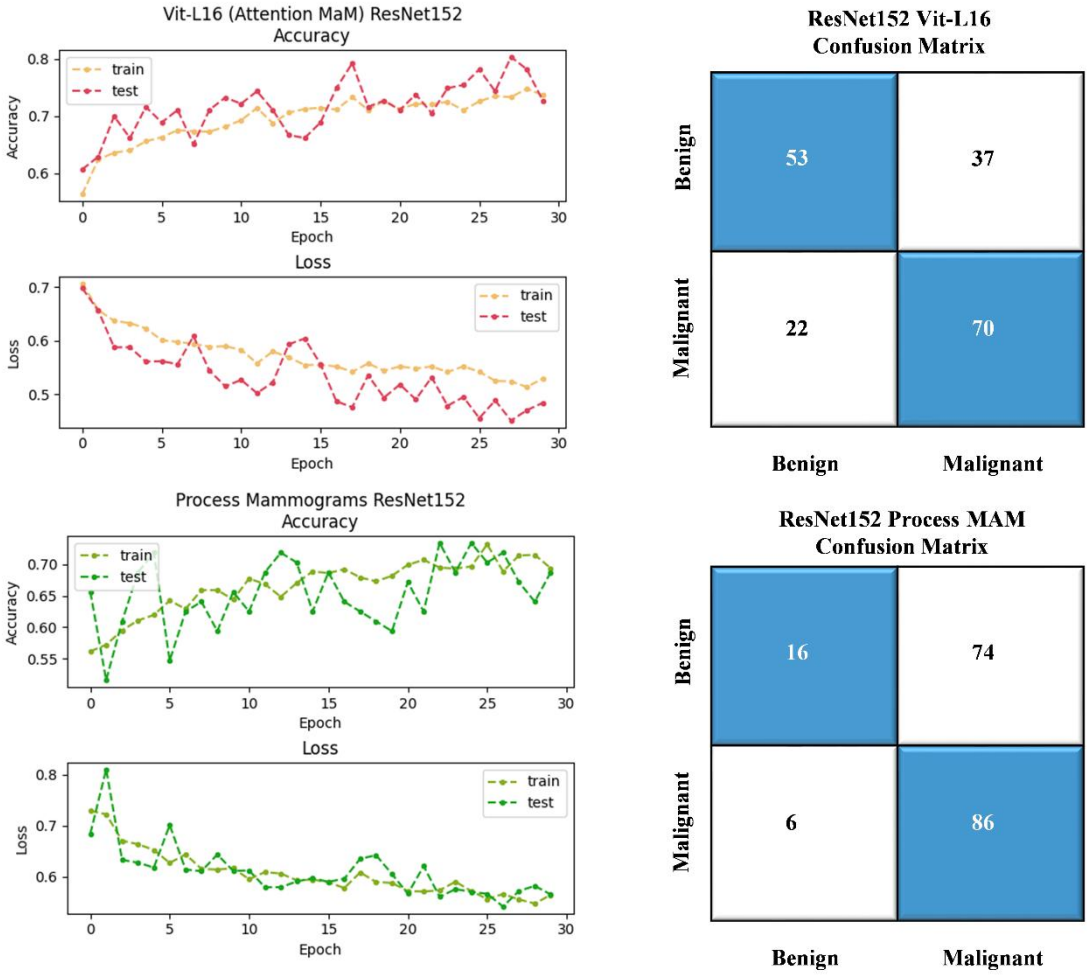

Figure 14. The training step ahead used pre-processed mammograms and attention VIT-L16 mammograms of ResNet152 measurement during 30 Epochs Accuracy, Error Loss, Confusion Matrices

Table 13. *WO and With Attention Mammograms ResNet152 Classification Report*

	With Attention		WO Attention	
Metrics	Benign	Malignant	Benign	Malignant
Sensitivity	0.18	0.93	0.74	0.68
Specificity	0.27	0.49	0.64	0.62
Precision	0.73	0.54	0.70	0.73
Recall	0.18	0.93	0.74	0.68
F-Score	0.29	0.68	0.72	0.71
	Accuracy	56.04	Accuracy	**71.42**

VGG16 and VGG19, based on the VGG architecture, exhibited competitive accuracy in classifying breast tumors. As illustrated in Figures 15 and 16, these models achieved notable accuracy values.

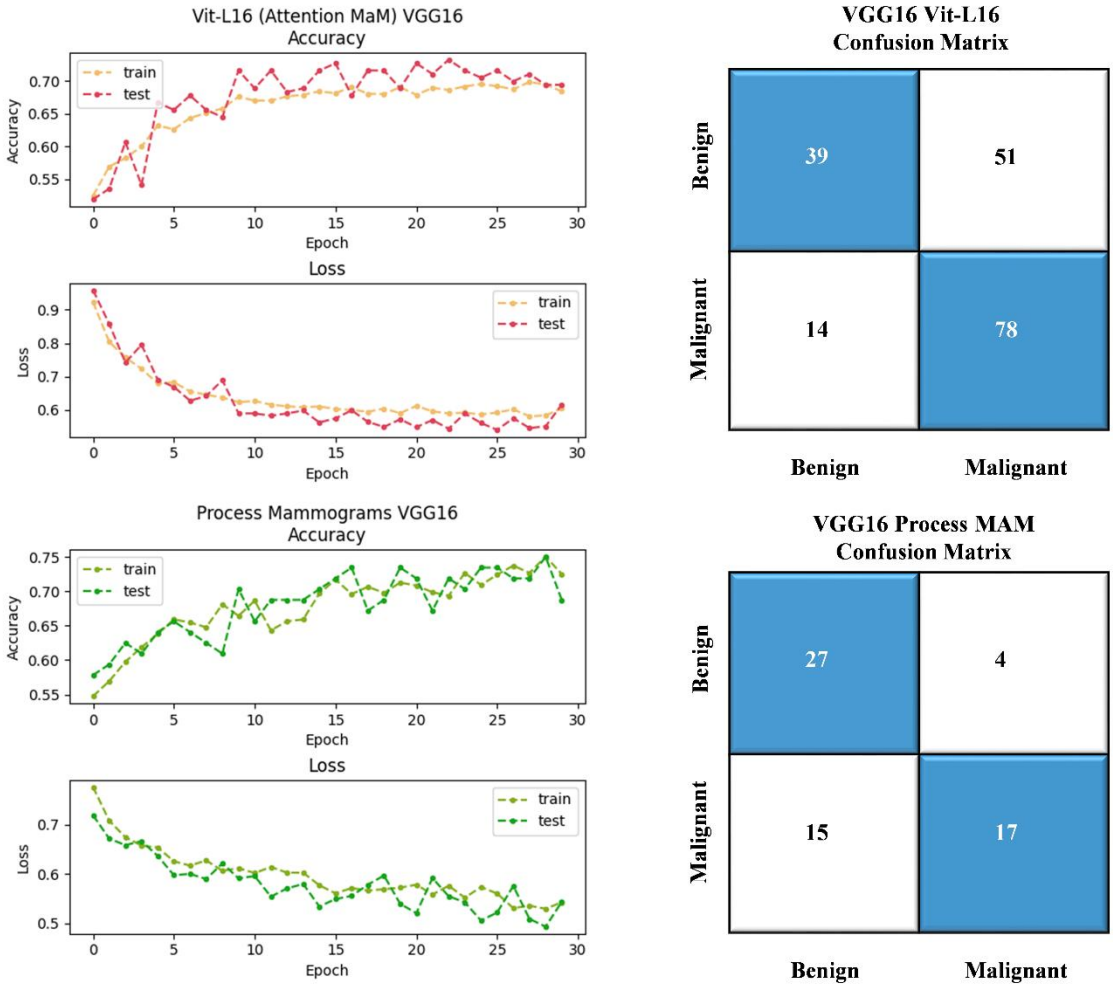

Figure 15. During 30 Epochs, the training Step forward using pre-process Mammograms and Attention VIT-L16 Mammograms of VGG16 measurement Accuracy, Error Loss, Confusion Matrices

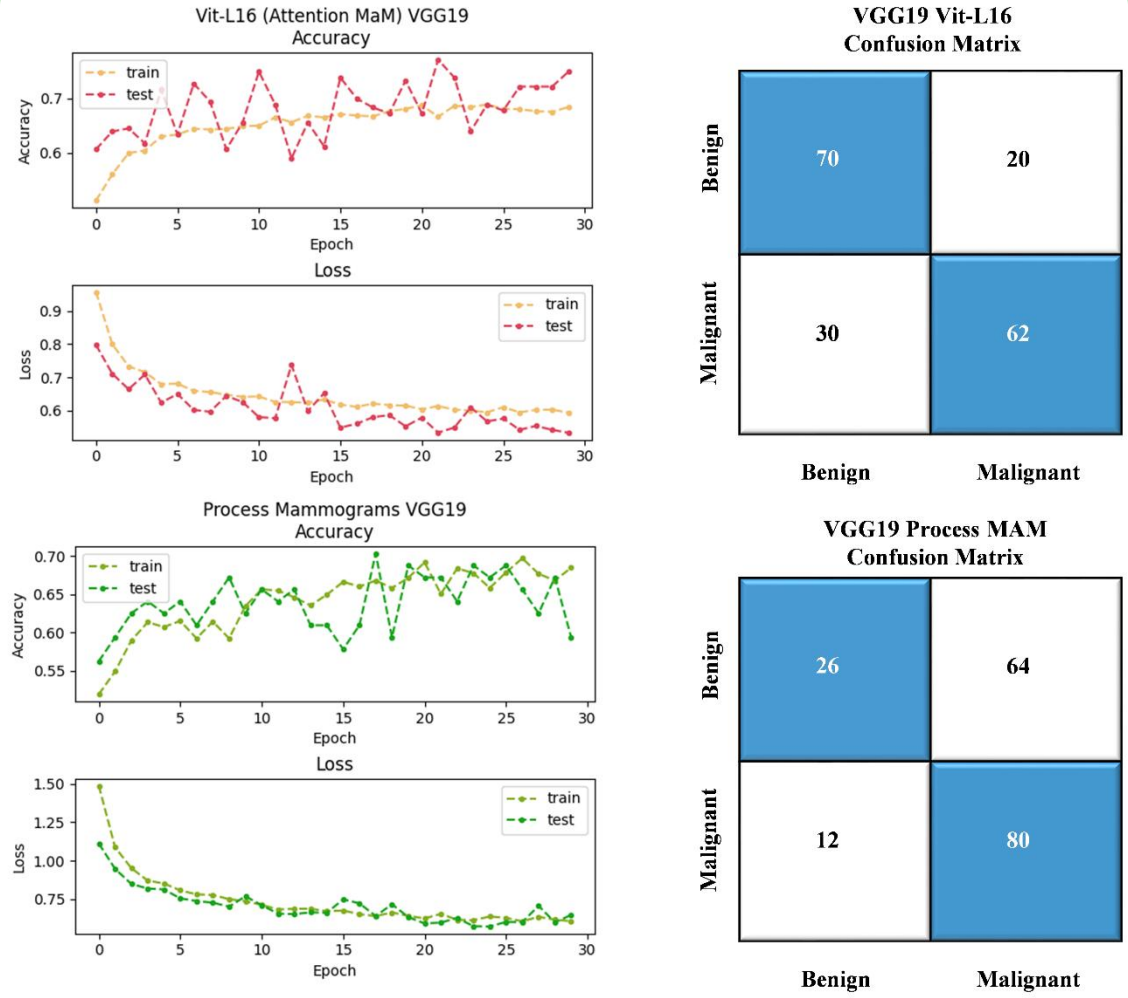

Figure 16. The training step ahead used pre-processed mammograms and attention VIT-L16 mammograms of VGG19 measurement during 30 Epochs Accuracy, Error Loss, Confusion Matrices

The classification reports in Tables 14 and 15 further support their performance. VGG models are known for their simplicity and uniform architecture, making them easy to understand and implement. However, their main drawback lies in their high number of parameters, which can make them computationally expensive and memory-intensive.

Table 14. *WO and With Attention Mammograms VGG16 Classification Report*

	WO Attention		With Attention	
Metrics	**Benign**	**Malignant**	**Benign**	**Malignant**
Sensitivity	0.87	0.53	0.43	0.85
Specificity	0.62	0.61	0.60	0.59
Precision	0.64	0.81	0.74	0.60
Recall	0.87	0.53	0.43	0.85
F-Score	0.74	0.64	0.55	0.71
	Accuracy	**69.84**	Accuracy	64.28

Table 15. *WO and With Attention Mammograms VGG19 Classification Report*

	WO Attention		With Attention	
Metrics	Benign	Malignant	Benign	Malignant
Sensitivity	0.29	0.87	0.78	0.67
Specificity	0.52	0.49	0.66	0.62
Precision	0.68	0.56	0.70	0.76

Recall	0.29	0.87	0.78	0.67
F-Score	0.41	0.68	0.74	0.71
	Accuracy	58.24	Accuracy	**72.52**

Xception, an extension of the Inception architecture, demonstrated impressive performance in terms of accuracy and loss during training. As shown in Figure 17, Xception achieved high accuracy values in classifying benign and malignant breast tumors. The confusion matrices in Figure 17 provide additional insights into the accuracy of each class. Xception leverages depth wise separable convolutions to reduce computational complexity while maintaining effective feature extraction capabilities.

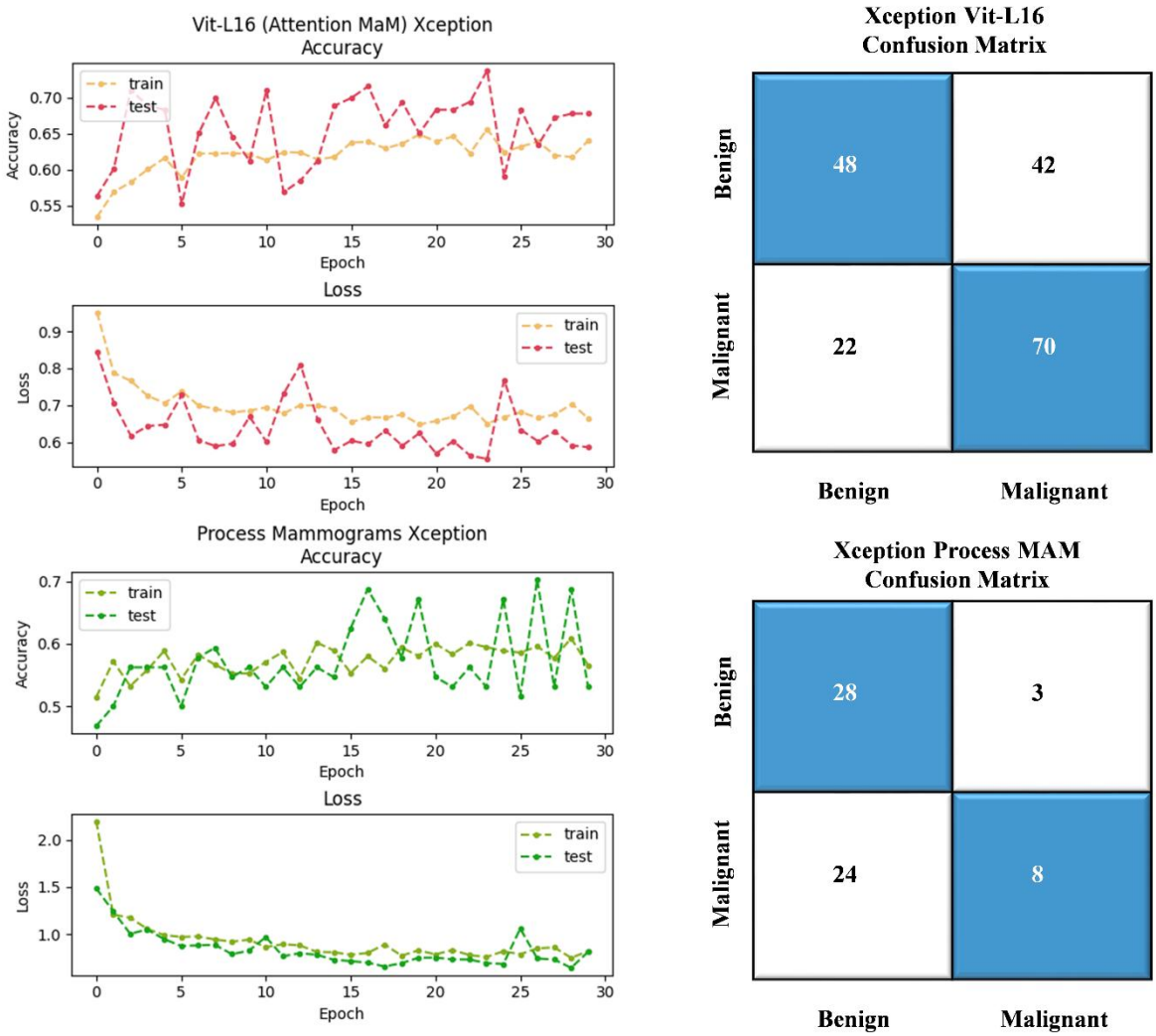

Figure 17. During 30 Epochs, the training Step forward using pre-process Mammograms and Attention VIT-L16 Mammograms of Xception measurement Accuracy, Error Loss, Confusion Matrices

However, Xception may require larger computational resources and longer training times compared to other models due to its increased complexity. Table 16 shows model performance with classification reports.

Table 16. *WO and With Attention Mammograms Xception Classification Report*

	WO Attention		With Attention	
Metrics	**Benign**	**Malignant**	**Benign**	**Malignant**
Sensitivity	0.90	0.25	0.70	0.76
Specificity	0.54	0.42	0.61	0.62

Precision	0.54	0.73	0.69	0.62
Recall	0.90	0.25	0.53	0.76
F-Score	0.67	0.37	0.60	0.69
	Accuracy	57.14	Accuracy	**64.83**

Overall, the evaluation of these fourteen deep convolutional neural network models for the binary classification of breast tumors on mammography images provides valuable insights into their performance and characteristics. While each model has its strengths and weaknesses, they collectively demonstrate the potential of deep learning in medical image analysis. Future research can further explore techniques to optimize and fine-tune these models to improve their performance, scalability, and efficiency in real-world clinical applications.

The effectiveness of fourteen deep convolution neural networks has been demonstrated in Table 17-18 both with and without attention. These tables include details on networks, including their total layers, slide input size, total parameters in millions, depth, and evaluation metrics. The ResNet50 model produced the best outcome 73.07% accuracy, and the VGG19 model came in second place when compared to other models with attention base transformer VIT-L16 mammograms. While the ResNet50 and ResNet101 deep transfer learning models have achieved the same highest accuracy 73.01% as compared to any other pre-trained model. Furthermore, as far as we are aware, this is the first comparison research to combine the advantages of two new technologies (Transformer, Deep Transfer Learning) to assess the performance of models with and without attention mechanisms.

Table 17. *With Attention Mammograms Performance of fourteen Deep Transfer Learning Model*

Models	Depth	Size	Parm (M)	Input	AC	RC	P	FS
AlexNet	8	227	61.0	224 x 224	50.68	0.51	0.26	0.34
DenseNet121	121	33	8.1	224 x 224	59.89	0.60	0.60	0.60
DensNet201	201	80	20.2	224 x 224	57.69	0.58	0.62	0.58
GoogleNet	22	27	7.0	224 x 224	50.54	0.51	0.51	0.49
MobileNet	53	16	4.3	224 x 224	56.04	0.56	0.56	0.56
NASNetMobile	Non-Linear	23	5.3	224 x 224	64.28	0.64	0.64	0.64
ResNet50	50	98	25.6	224 x 224	73.07	0.73	0.73	0.73
ResNet101	101	171	44.7	224 x 224	71.42	0.71	0.72	0.71
ResNet152	152	232	60.4	224 x 224	71.42	0.68	0.68	0.67
InceptionResNetv2	164	209	55.9	224 x 224	50.54	0.51	0.26	0.34
Inception	48	92	23.9	229 x 229	59.89	0.60	0.67	0.56
VGG16	16	528	138.4	224 x 224	64.28	0.65	0.68	0.65
VGG19	19	549	143.7	224 x 224	72.52	0.73	0.73	0.72
Xception	81	88	22.9	224 x 224	64.83	0.66	0.64	0.64

Table 18. *WO Attention Mammograms Performance of fourteen Deep Transfer Learning Model*

Models	Depth	Size	Parm (M)	Input	AC	RC	P	FS
AlexNet	8	227	61.0	224 x 224	50.79	0.51	0.26	0.34
DenseNet121	121	33	8.1	224 x 224	63.49	0.63	0.69	0.61
DensNet201	201	80	20.2	224 x 224	69.84	0.70	0.73	0.69
GoogleNet	22	27	7.0	224 x 224	66.66	0.67	0.68	0.66
MobileNet	53	16	4.3	224 x 224	63.49	0.63	0.64	0.63
NASNetMobile	NL	23	5.3	224 x 224	50.79	0.51	0.51	0.51
ResNet50	50	98	25.6	224 x 224	73.01	0.74	0.73	0.73
ResNet101	101	171	44.7	224 x 224	73.01	0.83	0.73	0.71
ResNet152	152	232	60.4	224 x 224	56.04	0.56	0.63	0.49
InceptionResNetv2	164	209	55.9	224 x 224	49.20	0.49	0.24	0.32
Inceptionv3	48	92	23.9	229 x 229	49.20	0.70	0.73	0.69

VGG16	16	528	138.4	224 x 224	69.84	0.70	0.73	0.69
VGG19	19	549	143.7	224 x 224	58.24	0.58	0.62	0.54
Xception	81	88	22.9	224 x 224	57.14	0.57	0.63	0.52

4. Discussion

Breast cancer remains a significant global health concern, being one of the leading causes of cancer-related deaths among women. Timely detection plays a crucial role in improving survival rates, and mammography has emerged as the most widely utilized screening method for breast cancer. With the increasing complexity of medical imaging, CAD systems have gained popularity as valuable tools to assist radiologists identify breast lesions. In this study, we extensively evaluated fourteen pre-trained deep learning models on the INbreast mammography dataset to assess their performance in breast cancer detection.

Our evaluation results highlight ResNet50 as the top-performing model for breast cancer detection. While the performance measures of the residual network and VGG models were comparable, ResNet50 demonstrated significantly higher accuracy in both attention mammograms and without attention mammograms. This superior performance can be attributed to ResNet50's deeper network structure, which enables it to address the vanishing gradient problem more effectively than VGG models. The problem of vanishing gradients arises when gradients become exceedingly small as they propagate through the layers of a deep neural network, hindering effective learning from the data. ResNet50 overcomes this challenge by incorporating residual connections that facilitate learning from residual error, resulting in improved performance on the INbreast mammography dataset.

In addition to ResNet50's remarkable performance, it is essential to highlight the limitations and drawbacks of the alternative models. While the residual network and VGG models demonstrated comparable performance, they exhibited challenges in handling the vanishing gradient problem, which can negatively impact learning capabilities. The vanishing gradient problem can lead to difficulties in capturing subtle patterns and features in mammograms, potentially reducing the accuracy of breast cancer detection. In contrast, ResNet50's ability to effectively mitigate the vanishing gradient problem through residual connections enables it to learn from the residual error, leading to enhanced performance.

Moreover, ResNet50 capacity to distinguish between benign and malignant tumors during the training process plays a vital role in reducing the risk of overfitting. Overfitting occurs when a model becomes overly complex and memorizes the training data rather than generalizing to new, unseen data. By accurately distinguishing between benign and malignant tumors, ResNet50 ensures that the model focuses on relevant features and avoids learning irrelevant characteristics, thereby reducing the risk of overfitting. In contrast, other models may face challenges in distinguishing between tumor types, potentially leading to overfitting issues and reduced accuracy.

The INbreast mammography dataset, consisting of 410 mammography slides, served as a robust benchmark for evaluating the models. Our evaluation results consistently demonstrated ResNet50's superiority across various metrics, including training validation curves, class-wise confusion matrix, and testing accuracy. The superior performance of ResNet50 can be attributed to its improved information flow from the input layer to the output layer, which facilitates smoother gradient propagation throughout the network. Furthermore, in our study, we also explored the performance of the Vit Transformer for computer vision applications. The transformer was used to identify attention in mammograms, and our findings revealed an increase in the accuracy of computer vision applications with the integration of the transformer. The attention mechanism employed by the transformer enables the model to focus on relevant regions of the mammogram image while minimizing the impact of irrelevant regions. This functionality proves to be particularly beneficial in mammography due to the high degree of variability in the appearance of breast tissue. However, it is important to note that the transformer encoder layer still requires further improvements to achieve even greater accuracy.

In conclusion, our evaluation results unequivocally establish ResNet50 as the optimal model for breast cancer detection on the

INbreast mammography dataset. Its better performance can be attributed to its effective handling of the vanishing gradient problem, its ability to differentiate between benign and malignant tumors, and its emphasis on rich features. The study also sheds light on the limitations of other models, highlighting challenges in addressing the vanishing gradient problem and distinguishing between tumor types. Additionally, we demonstrate the potential of attention mechanisms in computer vision applications, emphasizing the need for ongoing research and development to further improve accuracy. Finally, these findings contribute valuable insights to the advancement of deep learning models for breast cancer detection, ultimately enhancing the effectiveness of CAD systems in clinical practice.

5. Conclusion

This study evaluates fourteen deep convolutions to classify breast cancer, utilizing the INbreast dataset for performance. The models customize through the replacement of the final fully connected layer with our two layers architecture, improved performance was achieved. Results from Tables 15-16, with and without attention, indicated that the residual networks and VGG model consistently demonstrated higher overall accuracy. Notably, the Residual network outperformed other models, with the VGG model following closely behind. These findings provide a valuable starting point for developing more accurate and computationally efficient models. In addition, our research introduced a novel analysis of attention weight scores using Hierarchical Attention Networks (VIT-L16) combined with a deep dense transfer network. This analysis shed light on the factors influencing prediction outcomes, particularly in the context of success, representing the first exploration of its kind in breast cancer classification. The insights derived from this analysis have significant implications for researchers and practitioners, such as biomedical engineering and medical practitioner, seeking to make informed decisions and enhance the success rate of investments and campaigns. Furthermore, integrating attention mechanisms in the Transformer model, specifically in attention mammograms (Vit Transformer), increased accuracy in computer vision applications. However, further enhancements to the transformer encoder layer are required to maximize accuracy potential. This study establishes the superior performance of residual networks and the VGG model in breast cancer classification. Furthermore, the study highlights the potential of attention mechanisms in improving predictive outcomes and emphasizes the ongoing advancements needed in transformer models. Finally, these findings set the stage for future research aimed at developing more efficient and accurate models for breast cancer detection.

REFERENCES

1. Waqas, Muhammd, Muhammad Atif Tahir, and Rizwan Qureshi. "Ensemble-based instance relevance estimation in multiple-instance learning." In 2021 9th European workshop on visual information processing (EUVIP), pp. 1-6. IEEE, 2021.
2. Khan, M.A., Khan, S.U.R. & Lin, D. Shortening surgical time in high myopia treatment: a randomized controlled trial comparing non-OVD and OVD techniques in ICL implantation. BMC Ophthalmol 25, 303 (2025). https://doi.org/10.1186/s12886-025-04135-3
3. Shahzad, Inzamam, Jianquan Ouyang, and Saif Ur Rehman Khan. "FedVC-ADDiM: a federated learning framework for diagnosis of alzheimer disease using deep learning." Multimedia Systems 32, no. 3 (2026): 161. https://doi.org/10.1007/s00530-026-02229-6
4. Khan, S.U.R., Asif, S., Bilal, O. et al. Lead-cnn: lightweight enhanced dimension reduction convolutional neural network for brain tumor classification. Int. J. Mach. Learn. & Cyber. (2025). https://doi.org/10.1007/s13042-025-02637-6.
5. Khan, S.U.R., Zhao, M. & Li, Y. Detection of MRI brain tumor using residual skip block based modified MobileNet model. Cluster Comput 28, 248 (2025). https://doi.org/10.1007/s10586-024-04940-3
6. Hekmat, A., et al., Brain tumor diagnosis redefined: Leveraging image fusion for MRI enhancement classification. Biomedical Signal Processing and Control, 2025. 109: p. 108040.

7. Ishfaque, Muhammad, Saif Ur Rehman Khan, and Yulong Lou. "Digitizing Health Monitoring in Engineering Structures Using Deep Learning: A Novel Block Architecture for Concrete Crack Prediction in Surface and Sub-surface Dataset." Journal of Bionic Engineering (2026): 1-23.
8. Khan, Saif Ur Rehman, Asif Raza, Inzamam Shahzad, and Ghazanfar Ali. "Enhancing concrete and pavement crack prediction through hierarchical feature integration with VGG16 and triple classifier ensemble." In 2024 Horizons of Information Technology and Engineering (HITE), pp. 1-6. IEEE, 2024.
9. Waqas, Muhammad, Zeshan Khan, Shaheer Anjum, and Muhammad Atif Tahir. "Lung-Wise Tuberculosis Analysis and Automatic CT Report Generation with Hybrid Feature and Ensemble Learning." In CLEF (Working notes), pp. 1-10. 2020.
10. Khan, S. U. R., & Khan, Z. (2025). Detection of Abnormal Cardiac Rhythms Using Feature Fusion Technique with Heart Sound Spectrograms. Journal of Bionic Engineering, 1-20.
11. Waqas, Muhammad, Syed Umaid Ahmed, Muhammad Atif Tahir, Jia Wu, and Rizwan Qureshi. "Exploring multiple instance learning (MIL): A brief survey." Expert Systems with Applications 250 (2024): 123893.
12. Al-Khasawneh, Mahmoud Ahmad, Asif Raza, Saif Ur Rehman Khan, and Zia Khan. "Stock Market Trend Prediction Using Deep Learning Approach." Computational Economics (2024): 1-32
13. Khan, Muhammad Ahmed, Manqiang Peng, Ding Lin, and Saif Ur Rehman Khan. "Deep Learning Based Estimation of Blood Glucose Levels from Multidirectional Scleral Blood Vessel Imaging." arXiv preprint arXiv:2603.12715 (2026).
14. Khan, Saif Ur Rehman, Muhammad Nabeel Asim, Sebastian Vollmer, and Andreas Dengel. "FloraSyntropy-net: scalable deep learning with novel FloraSyntropy archive for large-scale plant disease diagnosis." Plant Methods (2026).
15. Ur Rehman Khan, Saif, Omair Bilal, Arash Hekmat, Inzamam Shahzad, and Asif Raza. "Advancing food safety: deep learning for accurate detection of bacterial contaminants." Memetic Computing 18, no. 1 (2026): 11.
16. Waqas, Muhammad, Muhammad Atif Tahir, Sumaya Al-Maadeed, Ahmed Bouridane, and Jia Wu. "Simultaneous instance pooling and bag representation selection approach for multiple-instance learning (MIL) using vision transformer." Neural Computing and Applications 36, no. 12 (2024): 6659-6680.
17. Hekmat, Arash, Omair Bilal, Zuping Zhang, Saif Ur Rehman Khan, and Sohaib Asif. "FRE-Net: A Fuzzy Richards Functions-Based Ensemble Network for Brain Tumor Detection." Journal of Bionic Engineering (2026): 1-23.
18. Mayumu, Nicanor, Xiaoheng Deng, Antoine Bagula, and Patrick Mukala. "V2X-JEPA: Self-Supervised Multi-Agent Joint Embedding Predictive Architecture for Robust Vehicle-to-Everything Perception." IEEE Internet of Things Journal (2026).
19. Waqas, Muhammad, Muhammad Atif Tahir, and Salman A. Khan. "Robust bag classification approach for multi-instance learning via subspace fuzzy clustering." Expert Systems with Applications 214 (2023): 119113.
20. Bilal, Omair, Arash Hekmat, Inzamam Shahzad, Asif Raza, and Saif Ur Rehman Khan. "Boosting Machine Learning Accuracy for Cardiac Disease Prediction: The Role of Advanced Feature Engineering and Model Optimization." The Review of Socionetwork Strategies (2025): 1-30.
21. Khan, Saif Ur Rehman, Muhammad Nabeel Asim, Sebastian Vollmer, and Andreas Dengel. "Temperature-driven robust disease detection in brain and gastrointestinal disorders via context-aware adaptive knowledge distillation." Biomedical Signal Processing and Control 112 (2026): 108671.
22. Khan, S. U. R., Asif, S., Zhao, M., Zou, W., Li, Y., & Xiao, C. (2026). ShallowMRI: A novel lightweight CNN with novel attention mechanism for Multi brain tumor classification in MRI images. Biomedical Signal Processing and Control, 111, 108425.
23. Khan, M. A., Khan, S. U. R., Rehman, H. U., Aladhadh, S., & Lin, D. (2025). Robust InceptionV3 with Novel EYENET Weights for Di-EYENET Ocular Surface Imaging

Dataset: Integrating Chain Foraging and Cyclone Aging Techniques. International Journal of Computational Intelligence Systems, 18(1), 1-26.

24. A. Raza, Salahuddin, S. Latif, and G. Ali, “Refine Security Control Protocols for Block chain in Textile Industry Supply Chain Management”, VFAST trans. softw. eng., vol. 14, no. 2, pp. 62–80, Apr. 2026.
25. Khan, U. S., & Khan, S. U. R. (2025). Ethics by Design: A Lifecycle Framework for Trustworthy AI in Medical Imaging From Transparent Data Governance to Clinically Validated Deployment. arXiv preprint arXiv:2507.04249.
26. Waqas, Muhammad, Muhammad Atif Tahir, and Rizwan Qureshi. "Deep Gaussian mixture model based instance relevance estimation for multiple instance learning applications." Applied intelligence 53, no. 9 (2023): 10310-10325.
27. Khan, S. U. R., Rehman, H. U., & Bilal, O. (2025). AI-powered cancer diagnosis: classifying viable (live) vs non-viable (dead) cells using transfer learning. Signal, Image and Video Processing, 19(15), 1326.
28. Khan, S. U. R., Asif, S., Zhao, M., Zou, W., Li, Y., & Xiao, C. (2026). ShallowMRI: A novel lightweight CNN with novel attention mechanism for Multi brain tumor classification in MRI images. Biomedical Signal Processing and Control, 111, 108425.
29. Waqas M, Bandyopadhyay R, Showkatian E, Muneer A, Zafar A, Alvarez FR, Marin MC, Li W, Jaffray D, Haymaker C, Heymach J. The Next Layer: Augmenting Foundation Models with Structure-Preserving and Attention-Guided Learning for Local Patches to Global Context Awareness in Computational Pathology. arXiv preprint arXiv:2508.19914. 2025 Aug 27.
30. Shahzad, Inzamam, Asif Raza, and Muhammad Waqas. "Medical Image Retrieval using Hybrid Features and Advanced Computational Intelligence Techniques." Spectrum of engineering sciences 3, no. 1 (2025): 22-65.
31. Bilal, O., Hekmat, A., Shahzad, I. et al. Boosting Machine Learning Accuracy for Cardiac Disease Prediction: The Role of Advanced Feature Engineering and Model Optimization. Rev Socionetwork Strat (2025). https://doi.org/10.1007/s12626-025-00190-w
32. Asif Raza, Inzamam Shahzad, Ghazanfar Ali, and Muhammad Hanif Soomro. "Use Transfer Learning VGG16, Inception, and Reset50 to Classify IoT Challenge in Security Domain via Dataset Bench Mark." Journal of Innovative Computing and Emerging Technologies 5, no. 1 (2025).
33. Khan, Z., Khan, S. U. R., Bilal, O., Raza, A., & Ali, G. (2025, February). Optimizing Cervical Lesion Detection Using Deep Learning with Particle Swarm Optimization. In 2025 6th International Conference on Advancements in Computational Sciences (ICACS) (pp. 1-7). IEEE.
34. Waqas, Muhammad, Zeshan Khan, Shaheer Anjum, and Muhammad Atif Tahir. "Lung-Wise Tuberculosis Analysis and Automatic CT Report Generation with Hybrid Feature and Ensemble Learning." In CLEF (Working notes), pp. 1-10. 2020.
35. Khan, S. U. R., Asif, S., Zhao, M., Zou, W., & Li, Y. (2025). Optimize brain tumor multiclass classification with manta ray foraging and improved residual block techniques. Multimedia Systems, 31(1), 1-27.
36. Asif Raza, Salahuddin, Ghazanfar Ali, Muhammad Hanif Soomro, Saima Batool, "Analyzing the Impact of Artificial Intelligence on Shaping Consumer Demand in E-Commerce: A Critical Review", International Journal of Information Engineering and Electronic Business(IJIEEB), Vol.17, No.5, pp. 42-61, 2025. DOI:10.5815/ijieeb.2025.05.04
37. Khan, M. A., Khan, S. U. R., Rehman, H. U., Aladhadh, S., & Lin, D. (2025). Robust InceptionV3 with Novel EYENET Weights for Di-EYENET Ocular Surface Imaging Dataset: Integrating Chain Foraging and Cyclone Aging Techniques. International Journal of Computational Intelligence Systems, 18(1), 204.
38. Raza, Asif, Inzamam Shahzad, Muhammad Salahuddin, and Sadia Latif. "Satellite Imagery Employed to Analyze the Extent of Urban Land Transformation in The Punjab District of Pakistan." Journal of Palestine Ahliya University for Research and Studies 4, no. 2 (2025): 17-36.

39. Khan, S. U. R., Asif, S., Zhao, M., Zou, W., Li, Y., & Li, X. (2025). Optimized deep learning model for comprehensive medical image analysis across multiple modalities. Neurocomputing, 619, 129182.
40. Khan, Saif Ur Rehman, Asif Raza, Inzamam Shahzad, and Shehzad Khan. "Subcellular Structures Classification in Fluorescence Microscopic Images." In International Conference on Computing & Emerging Technologies, pp. 271-286. Cham: Springer Nature Switzerland, 2023.
41. Maqsood, H., & Khan, S. U. R. (2025). MeD-3D: A Multimodal Deep Learning Framework for Precise Recurrence Prediction in Clear Cell Renal Cell Carcinoma (ccRCC). arXiv preprint arXiv:2507.07839.
42. Raza, Asif, Salahuddin, & Inzamam Shahzad. (2024). Residual Learning Model-Based Classification of COVID-19 Using Chest Radiographs. Spectrum of Engineering Sciences, 2(3), 367–396.
43. Khan, S. U. R. (2025). Multi-level feature fusion network for kidney disease detection. Computers in Biology and Medicine, 191, 110214.
44. Salahuddin, Syed Shahid Abbas, Prince Hamza Shafique, Abdul Manan Razzaq, & Mohsin Ikhlaq. (2024). Enhancing Reliability and Sustainability of Green Communication in Next-Generation Wireless Systems through Energy Harvesting. Journal of Computing & Biomedical Informatics.
45. S. U. R. Khan, A. Raza, I. Shahzad and G. Ali, "Enhancing Concrete and Pavement Crack Prediction through Hierarchical Feature Integration with VGG16 and Triple Classifier Ensemble," 2024 Horizons of Information Technology and Engineering (HITE), Lahore, Pakistan, 2024, pp. 1-6.
46. Mahmood, F., Abbas, K., Raza, A., Khan,M.A., & Khan, P.W. (2019). Three Dimensional Agricultural Land Modeling using Unmanned Aerial System (UAS). International Journal of Advanced Computer Science and Applications (IJACSA) [p-ISSN : 2158-107X, e-ISSN : 2156-5570], 10(1).
47. HUSSAIN, S., Raza, A., MEERAN, M. T., IJAZ, H. M., & JAMALI, S. (2020). Domain Ontology Based Similarity and Analysis in Higher Education. IEEEP New Horizons Journal, 102(1), 11-16.
48. Hekmat, A., Zuping, Z., Bilal, O., & Khan, S. U. R. (2025). Differential evolution-driven optimized ensemble network for brain tumor detection. International Journal of Machine Learning and Cybernetics, 1-26.
49. Raza, A., & Meeran, M. T. (2019). Routine of Encryption in Cognitive Radio Network. Mehran University Research Journal of Engineering and Technology [p-ISSN: 0254-7821, e-ISSN: 2413-7219], 38(3), 609-618.
50. Meeran, M. T., Raza, A., & Din, M. (2018). Advancement in GSM Network to Access Cloud Services. Pakistan Journal of Engineering, Technology & Science [ISSN: 2224-2333], 7(1).
51. Bilal, O., Hekmat, A., & Khan, S. U. R. (2025). Automated cervical cancer cell diagnosis via grid search-optimized multi-CNN ensemble networks. Network Modeling Analysis in Health Informatics and Bioinformatics, 14(1), 67.
52. Raza, Asif , Soomro, M. H., Shahzad, I., & Batool, S. (2024). Abstractive Text Summarization for Urdu Language. Journal of Computing & Biomedical Informatics, 7(02).
53. M. Wajid, M. K. Abid, A. Asif Raza, M. Haroon, and A. Q. Mudasar, “Flood Prediction System Using IOT & Artificial Neural Network”, VFAST trans. softw. eng., vol. 12, no. 1, pp. 210–224, Mar. 2024.
54. N. Mayumu, D. Xiaoheng, P. Mukala, S. U. R. Khan and M. U. Saeed, "Omni-V2X: A Vision-Language Model for Actionable Insights in Vehicle-to-Everything Systems," 2025 International Joint Conference on Neural Networks (IJCNN), Rome, Italy, 2025, pp. 1-8, doi: 10.1109/IJCNN64981.2025.11228491.
55. Waqas, M., Vierra, C., Kaplan, D. L., & Othman, S. (2019). Feasibility of low field MRI and proteomics for the analysis of tissue engineered bone. Biomedical Physics & Engineering Express, 5(2), 025037.
56. Khan, S. R., Asif Raza, Inzamam Shahzad, & Hafiz Muhammad Ijaz. (2024). Deep transfer CNNs models performance evaluation using unbalanced histopathological breast cancer dataset.

Lahore Garrison University Research Journal of Computer Science and Information Technology, 8(1).
57. M. Waqas, Z. Khan, S. U. Ahmed and Asif. Raza, "MIL-Mixer: A Robust Bag Encoding Strategy for Multiple Instance Learning (MIL) using MLP-Mixer," 2023 18th International Conference on Emerging Technologies (ICET), Peshawar, Pakistan, 2023, pp. 22-26.
58. S. ur R. Khan, Asif. Raza, Muhammad Tanveer Meeran, and U. Bilhaj, "Enhancing Breast Cancer Detection through Thermal Imaging and Customized 2D CNN Classifiers", VFAST trans. softw. eng., vol. 11, no. 4, pp. 80–92, Dec. 2023.
59. Yang, H., Khan, S. U. R., Bilal, O., Chen, C., & Zhao, M. (2025). CEOE-Net: Chaotic Evolution Algorithm-Based Optimized Ensemble Framework Enhanced with Dual-Attention for Alzheimer's Diagnosis. *Computer Modeling in Engineering & Sciences*, *145*(2), 2401.
60. Chomba, B., Mukala, P., Mayumu, N., & Khan, S. U. R. (2025). DynaKG: Dynamic Knowledge Graph Attention With Learnable Temporal Decay for Recommendation. *IEEE Access*, *13*, 216956-216970.
61. O. Bilal, Asif Raza, S. ur R. Khan, and Ghazanfar Ali, "A Contemporary Secure Microservices Discovery Architecture with Service Tags for Smart City Infrastructures ", VFAST trans. softw. eng., vol. 12, no. 1, pp. 79–92, Mar. 2024
62. S. U. R. Khan, A. Raza, I. Shahzad and G. Ali, "Enhancing Concrete and Pavement Crack Prediction through Hierarchical Feature Integration with VGG16 and Triple Classifier Ensemble," 2024 Horizons of Information Technology and Engineering (HITE), Lahore, Pakistan, 2024, pp. 1-6, doi: 10.1109/HITE63532.2024.10777242.
63. Ijaz, M., Khan, S. U. R., Rehman, A. U., Zaib, A., Vollmer, S., Dengel, A., & Asim, M. N. (2026). SolarFCD: A Large-Scale Dataset and Benchmark for Solar Fault Classification in Photovoltaic Systems. arXiv preprint arXiv:2604.23662.
64. Ishfaque, M., Khan, S. U. R., & Lou, Y. L. (2026). Towards efficient dam inspection: crack detection via chirplet transform feature and a pruned VGG16 architecture. *Memetic Computing*, *18*(1), 9.
65. S. Ur Rehman Khan, O. Bilal, S. Mistry, N. Deb, M. Mahmud and M. Bhuyan, "KDLight: A Lightweight Knowledge Distillation Framework for Medical Image Classification," *2025 International Joint Conference on Neural Networks (IJCNN)*, Rome, Italy, 2025, pp. 1-8, doi: 10.1109/IJCNN64981.2025.11228615.
66. O. Bilal, S. Ur Rehman Khan, S. Mistry, N. Deb, M. Mahmud and M. Bhuyan, "Towards Efficient Pruning and Multi-Scale Feature Transformations to Uncover Medical Diseases," *2025 International Joint Conference on Neural Networks (IJCNN)*, Rome, Italy, 2025, pp. 1-8, doi: 10.1109/IJCNN64981.2025.11229047.
67. Waqar, I. A., Zaib, A., Ahmed, S., Vollmer, S., Dengel, A., & Asim, M. N. (2026). MS-SSE-Net: A Multi-Scale Spatial Squeeze-and-Excitation Network for Structural Damage Detection in Civil and Geotechnical Engineering. arXiv preprint arXiv:2604.14711.
68. Khan, S. R., Raza, A., Waqas, M., & Raphay Zia, M. A. (2024). Efficient and Accurate Image Classification via Spatial Pyramid Matching and SURF Sparse Coding. Lahore Garrison University Research Journal of Computer Science and Information Technology, 7(4).
69. Shahzad, Inzamam, Asif Raza, Hasaan Maqsood, Saif Ur Rehman Khan, and Ghazanfar Ali. "Towards Robust Breast Cancer Diagnosis: A Hybrid Deep Learning Ensemble Framework." In 2025 Horizons of Information Technology and Engineering (HITE), pp. 1-6. IEEE, 2025.
70. Ijaz, M., Khan, S. U. R., Rehman, A. U., Asif, T., Vollmer, S., Dengel, A., & Asim, M. N. (2026). GlobalWasteData: A Large-Scale, Integrated Dataset for Robust Waste Classification and Environmental Monitoring. arXiv preprint arXiv:2602.07463.
71. Bilal, O., Hekmat, A., Khan, S. U. R., Raza, A., & Ali, G. (2025, December). MS-STO-Net: A Multi-Scale State Transition Optimization-Based Ensemble Network for Accurate White Blood Cell Classification.

In *2025 27th International Multitopic Conference (INMIC)* (pp. 1-6). IEEE.

72. N. Mayumu, X. Deng, A. Bagula, S. u. R. Khan and P. Mukala, "V2X-JEPA: Self-Supervised Multi-Agent Joint Embedding Predictive Architecture for Robust Vehicle-to-Everything Perception," in *IEEE Internet of Things Journal*, doi: 10.1109/JIOT.2026.3660030.
73. Khan, S. U. R., Asim, M. N., Vollmer, S., & Dengel, A. (2025). Dynamic Weight Adjustment for Knowledge Distillation: Leveraging Vision Transformer for High-Accuracy Lung Cancer Detection and Real-Time Deployment. arXiv preprint arXiv:2510.20438.
74. Waqas, M., & Khan, M. A. (2018). JSOPT: A framework for optimization of JavaScript on web browsers. Mehran University Research Journal of Engineering & Technology, 37(1), 95-104.