SSRG International Journal of Electronics and Communication Engineering Volume 12 Issue 1, 257-270, January 2025
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V12I1P120

Original Article

Enhanced Breast Cancer Detection and Classification Using CNN-EfficientNet B3: A Deep Learning Approach with Segmentation and Feature Extraction from MIAS Dataset

K. Punithavathi[1], G. Yamuna[2]

[1]*Department of Electronics and Communication Engineering, Faculty of Engineering and Technology, Annamalai University, Tamil Nadu, India.*
[2]*Department of Electronics and Communication Engineering, Faculty of Engineering and Technology, Annamalai University, Tamil Nadu, India.*

[1]*Corresponding Author : punithavathikrishnamoorthy@gmail.com*

Received: 24 November 2024 Revised: 29 December 2024 Accepted: 16 January 2025 Published: 30 January 2025

***Abstract** - Breast Cancer (BC), characterized by the uncontrolled proliferation of breast cells, remains the most prevalent and life-threatening cancer affecting women. Early detection significantly increases the chances of survival by enabling timely medical intervention. Numerous methods have been proposed for breast cancer detection; however, limitations in diagnostic accuracy and efficiency persist. To address these challenges, this study introduces a robust deep-learning framework leveraging fine-tuned EfficientNet-B3 for the detection and classification of breast tumors. The methodology employs segmentation techniques to accurately delineate the affected breast tumor regions, reducing training complexity while enhancing classification precision. The model was trained and evaluated using the Mammographic Image Analysis Society (MIAS) dataset, incorporating critical preprocessing steps to optimize image quality and feature extraction. Fine-tuning of EfficientNet-B3 was carried out to adapt the pre-trained network to mammogram-specific features, with hyperparameters optimized for this domain. Performance was assessed using three primary evaluation metrics: accuracy, specificity, and sensitivity. Experimental results demonstrate that the proposed CNN-EfficientNet-B3 framework achieves superior performance compared to conventional approaches, with specificity, accuracy, and sensitivity rates of 97.12%, 96.50%, and 96.43%, respectively. These findings highlight the potential of the proposed method to significantly enhance breast cancer detection and classification, paving the way for more effective clinical applications.*

***Keywords** - Convolutional Neural Networks, Deep Learning, Breast Cancer, EfficientNet, Mammogram.*

1. Introduction

Cancer is caused by genetic mutations that lead to abnormal changes in the genes responsible for regulating cell development and maintaining their health. These genes, found in the cell's nucleus, are the "control room" of each cell, dictating processes such as cell division, growth, and death. Under normal circumstances, the body follows an orderly process where healthy new cells replace older, dying ones. However, mutations can occur in specific genes, leading them to either activate or deactivate, causing a disturbance in the balance. Mutated cells acquire the ability to divide uncontrollably, generating more mutated cells that eventually form a tumor. Breast cancer, one of the most prevalent types of cancer among women, is the result of such mutations. It is a major cause of death globally, making early detection and intervention critical to reducing mortality rates [1]. Breast cancer detection has always been a challenge due to the complexity of the disease and the fact that its underlying causes remain largely unidentified. Despite this, early detection significantly improves the chances of successful treatment and complete recovery. Mammography has become the most commonly used tool for detecting both benign and malignant tumors in the breast. As technology has advanced, various imaging techniques such as Magnetic Resonance Imaging (MRI), ultrasound, and thermal imaging have been developed to further assist in the early identification of breast cancer. These methods have contributed to lowering recurrence rates and mortality, but challenges remain in detecting small anomalies in the early stages. Expert radiologists, despite their expertise, continue to miss a substantial proportion of anomalies in X-ray images due to the overwhelming number of images they must review. This issue underscores the importance of more efficient diagnostic methods that can assist radiologists in identifying early-stage breast cancer [2].

Recent advancements in mammography techniques have contributed to better detection rates. Digital mammography, Computer-Aided Detection (CAD), and breast tomosynthesis are some of the notable developments. These techniques have shown promise in enhancing the accuracy of cancer detection and minimizing the chances of missing early-stage tumors. However, they are not without their limitations, such as the difficulty in interpreting complex images, high false positive rates, and the need for human intervention. To overcome these challenges, machine learning and deep learning approaches have gained traction in recent years, promising significant improvements in the detection and classification of breast cancer [3].

1.1. Early Detection and Importance of Mammography

Mammography remains the gold standard for detecting breast cancer at an early stage. It involves taking X-ray images of the breast tissue to identify abnormalities, such as lumps or calcifications, which could indicate the presence of cancerous cells. Early detection is critical because it allows for the administration of treatments before the cancer spreads or grows into a more aggressive form, which could otherwise lead to the need for invasive procedures like mastectomy.

The likelihood of successful treatment for breast cancer increases when it is detected in its earliest stages. This is especially important because early-stage breast cancer often has a high survival rate when treated promptly. As the disease progresses, however, the chances of survival decrease significantly, making early detection a key factor in improving outcomes. Research has shown that the survival rate for breast cancer patients can be as high as 99% when detected at stage 1, compared to only 27% for those diagnosed at stage 4 [5]. mammography allows radiologists to detect breast cancer masses, both benign and malignant before they become palpable or cause noticeable symptoms.

This ability to detect presymptomatic cancer cells is crucial in preventing the disease from progressing into a more serious stage. However, despite the significant role of mammography in early cancer detection, there are limitations. One of the major challenges is the potential for missed diagnoses. Radiologists often need to examine a large number of images, which can lead to fatigue and, consequently, errors in interpretation. Additionally, subtle lesions and small tumors may not be easily visible in standard mammograms, making them difficult to detect. In some cases, lesions that are clearly visible to human eyes might still be overlooked due to image quality or the radiologist's experience. Recent advancements in mammography technology, such as digital mammography and Computer-Aided Detection (CAD), aim to address these challenges. These technologies offer improved image clarity, better visualization of dense breast tissue, and automated systems to help identify abnormalities. Digital mammography uses digital detectors to capture breast images and store them in computer files, allowing for more precise measurements and faster image retrieval. Computer-Aided Detection (CAD) systems analyze mammograms using algorithms to highlight areas that may need closer attention, providing additional support for radiologists.

Breast tomosynthesis, another development, uses multiple X-ray images to create a 3D reconstruction of the breast tissue, offering a more detailed view and helping to improve tumor detection [6]. Despite these advancements, breast cancer detection remains a complex task. Radiologists may still struggle to differentiate between benign and malignant lesions, especially when the tumor is small or located in dense tissue. Furthermore, the growing amount of medical images has made it increasingly difficult for radiologists to maintain high levels of accuracy. This problem, known as "image overload," highlights the need for more effective tools to assist in the detection process.

1.2. Limitations of Current Techniques

While mammography and its advanced variants, such as digital mammography and breast tomosynthesis, have undoubtedly improved the early detection of breast cancer, several limitations persist. The first challenge is the issue of false positives and false negatives. False positives occur when a benign tumor is mistakenly identified as malignant, leading to unnecessary biopsies or treatments. On the other hand, false negatives occur when malignant tumors are missed, which can delay treatment and reduce the chances of survival. These errors are particularly concerning in the early stages of cancer, where small tumors may be difficult to detect and easily overlooked. Another limitation is the variability in image quality.

Factors such as the patient's breast density, the imaging equipment used, and the radiologist's experience can all affect the accuracy of mammograms. For example, dense breast tissue can obscure the visibility of tumors, making them harder to identify. Moreover, the interpretation of mammographic images is subjective and dependent on the radiologist's expertise. This subjectivity can lead to inconsistencies in diagnoses, with some tumors being missed or misclassified.

In addition, there is the issue of limited access to mammography, particularly in rural areas or low-income regions. Although mammography is widely available in developed countries, there are still significant gaps in access in many parts of the world. This disparity in access means that many women in underserved regions may not receive timely screenings, which could lead to delayed diagnoses and poorer outcomes. Lastly, the high cost of mammography, along with the need for skilled radiologists to interpret the images, adds to the financial burden of healthcare systems. These challenges highlight the need for new, more efficient approaches to breast cancer detection that can reduce costs, improve accuracy, and overcome the limitations of current methods.

1.3. The Role of Deep Learning in Breast Cancer Detection

To address these challenges, researchers have increasingly turned to machine learning and deep learning techniques. These approaches offer the potential to enhance the accuracy of breast cancer detection by automating the process of image analysis. Deep learning, in particular, has shown great promise in this regard. By training neural networks on large datasets of mammographic images, deep learning models can learn to recognize complex patterns and features that may not be immediately apparent to human observers.

Deep learning models, such as Convolutional Neural Networks (CNNs), have been applied to breast cancer detection with significant success. These models are designed to automatically extract features from images and use them to classify the presence of cancerous cells. In some studies, CNNs have been shown to outperform traditional image processing techniques in terms of accuracy and sensitivity. The advantage of deep learning lies in its ability to learn from vast amounts of data, enabling the model to generalize and improve its performance over time.

One of the major benefits of using deep learning for breast cancer detection is the potential for reducing human error. By automating the process of analyzing mammograms, deep-learning models can help radiologists identify anomalies more quickly and accurately. These models can also reduce the burden on radiologists, allowing them to focus on more complex cases and improving efficiency.

Recent studies have shown that deep learning models can achieve high accuracy in detecting breast cancer, with some models achieving sensitivity and specificity rates comparable to or even exceeding those of human radiologists. These models are particularly effective in detecting small tumors and subtle anomalies that might be missed in traditional mammograms. Furthermore, deep learning models can be trained to differentiate between benign and malignant tumors, helping to reduce the number of false positives and false negatives.

Despite these promising results, challenges remain in implementing deep learning models in clinical practice. One of the main issues is the need for large, labeled datasets to train the models. While datasets such as the Mammographic Image Analysis Society (MIAS) dataset and the Digital Database for Screening Mammography (DDSM) provide valuable resources, there is still a need for more diverse and representative datasets that can improve the generalizability of the models. Additionally, deep learning models require significant computational power and resources, which may be a barrier to their widespread adoption in resource-limited settings.

2. Related Work

Recent advancements in deep learning have significantly enhanced breast cancer detection, enabling more accurate and timely diagnoses. In the context of mammography image analysis, various deep learning models have been proposed to improve the detection and classification of malignant and benign tumours. These advancements have leveraged a range of architectures, from traditional Convolutional Neural Networks (CNNs) to more complex models like EfficientNet, ResNet, and DenseNet.

One of the pioneering approaches involved CNN-based models for feature extraction and classification. LeNet and AlexNet, early CNN architectures, laid the foundation for subsequent models but were limited by their relatively shallow architectures and inability to capture fine-grained features essential for accurate tumour detection. More recent architectures, such as VGGNet, ResNet, and DenseNet, have improved on these by introducing deeper layers and skip connections, enabling better feature extraction from high-resolution images.

These models have been shown to achieve superior performance in classifying breast tumours but still face challenges in generalizing to diverse datasets.EfficientNet, as used in the current study, represents a significant breakthrough by using a compound scaling method that uniformly scales all dimensions of the network (depth, width, and resolution) to achieve higher accuracy with fewer parameters.

Studies have demonstrated that EfficientNet models outperform traditional CNN architectures, offering a more efficient solution for mammography image classification. Specifically, EfficientNet B3, used in this research, has been shown to deliver a balance between model size and performance, achieving state-of-the-art results on several medical imaging benchmarks.

Another notable approach has been the integration of ensemble learning techniques, where multiple models are combined to improve predictive accuracy. Methods like stacking or boosting have been explored to combine the strengths of different models, leading to improved sensitivity and specificity in breast cancer detection. Additionally, transfer learning techniques have been widely adopted, where pre-trained models on large datasets, such as ImageNet, are fine-tuned on smaller medical datasets, allowing for better generalization and reduced training time. Recent studies have also explored the use of advanced architectures like U-Net for segmentation tasks, which focus on delineating tumour boundaries in mammographic images. These methods are particularly effective in enhancing the model's interpretability and aiding radiologists in identifying tumours more accurately. Combining segmentation with classification in a multi-task learning framework has shown promising results in improving both detection and localization.

Despite these advancements, challenges remain, particularly in handling class imbalance, overfitting, and the availability of high-quality annotated data. Several works have addressed these issues through data augmentation, synthetic data generation, and the use of advanced regularization techniques like dropout and batch normalization. The integration of clinical metadata, including patient age, breast density, and family history, is also being explored to enhance model accuracy further.

In comparison to these approaches, the model proposed in this study CNN-EfficientNet B3 offers significant improvements in terms of both accuracy and computational efficiency. By leveraging a novel deep feature extraction strategy, using HOG for haze removal and a serial-based technique for feature selection, this model outperforms traditional methods and has shown excellent precision, sensitivity, and specificity. Moreover, the use of enhanced data through augmentation techniques ensures better generalization and robustness, making this approach well-suited for real-world clinical applications.

Breast cancer is the term for a malignant tumor that develops from breast cells. “Breast cancer usually begins in the lobules, the ducts or breastmilk-producing glands, which are the paths that transport milk flowing to the nipple from the lobules. Breast cancer is less common to arise from the histological tissues of the breast, which are made up of its fatty and fibrous connective tissues.

To make a diagnosis in the context of radiographic image interpretation, image data must be assessed and organised. The challenges related to these operations can include varying breast parenchyma with structural noise caused by thick tissue masking” [6]. These factors can make cancer lesions unnoticed or hidden. Manual diagnosis requires a number of subjective judgements, more diversity between and among observers, and the potential for serious errors [7].

A tumour may be benign (not harmful to health) or malignant (possibly harmful to health). “Benign tumors are not considered to be cancerous since they grow slowly, include cells that resemble healthy tissue, don't move to other parts of the body and don't infect nearby tissues. Cancerous tumours are called as malignant tumours. If malignant cells are not contained, they may eventually spread to other body organs. Clinical experts have fervently fought against tolerating "false alarms" [8]. Consequently, the creation of the Computer-Aided Diagnostic system (CAD) adds value by assisting the radiologist in reducing the number of false-negative and false-positive instances.

Additionally, while ground truth exists, the radiologist's expertise continues to serve as the benchmark and the deciding factor. The classifications for radiologists show that they are in the training stage [9, 10]. Radius (the average distance between the perimeter's points and the center), texture (the values of the grayscale standard deviation), the perimeter, smoothness, region, concavity, compactness, symmetry or fractal dimension, and concave points are examples of the physical characteristics that were measured in order to identify two classes of tumors and establish which class each tumor in new samples belongs to”.

Gaining an appropriate categorization, meanwhile, is still difficult because of several imaging problems and variances in the tumour regions. In order to detect and categorise medical infections throughout the past 10 years, AI has been very important for breast cancer [12]. The right course of therapy must be chosen in order to preserve and improve quality of life, and this requires accurately determining whether a tumor is cancerous or benign. The two primary genetic causes of breast cancer are damaged DNA and heritage.

However, other risk factors can be environmental or related to a person's way of life, such as alcohol consumption (studies show that women who drink three drinks per day have a 1.5 times higher risk of being affected), obesity, hormonal treatments (an increased level of oestrogen due to hormone replacement therapy or birth control pills can link to breast cancer), and a sedentary lifestyle without regular exercise. Additional factors could include delivery later in life or insufficient nursing.

This is founded on a number of intermediary procedures, such as preprocessing. of source images. Thanks to the CAD system, feature learning and feature extraction then, feature selection and feature reduction, and also classification were possible. The biggest triggers, however, are ignorance of available treatments and screening methods. The paucity of skilled radiologists and diagnostic resources, as well as the delay in providing critical care, are major issues. In an effort to aid this expanding cause, the plan is to develop deep learning models that can recognise worrisome lesions and, therefore, provide quick and reliable diagnoses. During the preliminary phase, the researcher strives to generate images of exceptional quality and eliminate any noise. The purpose of preprocessing is to make the cancer zone appear more prominently so that it may later be precisely identified as a Region of Interest (ROI) [13]. Table 1 shows the review of histopathology image detection and classification based on the deep learning technique used for breast cancer.

3. Proposed Methodology

3.1. Major Challenges and Contributions

Medical image processing, particularly the classification of mammogram images using deep learning, encounters several significant challenges. One of the primary issues is the limited availability of mammography image datasets. Deep learning models require extensive training data to achieve robust generalization, but the scarcity of datasets in this domain poses a significant hurdle. Additionally, these models demand a deeper understanding of the classification objective, which becomes challenging given the subtle and complex nature of emerging features in mammographic images.

Another critical challenge arises from the presence of redundant or irrelevant features in the dataset. These redundant features not only increase computational complexity and training time but also contribute to a higher false-negative rate, which is detrimental in medical diagnostics where accurate detection is paramount. In response to these challenges, this research makes several key contributions.

It introduces a method for enhancing mammography image clarity using haze removal techniques, improving feature representation for deep learning. The study also incorporates a novel data augmentation strategy to address the issue of limited datasets, ensuring the CNN-EfficientNet-B3 model learns effectively from diverse and enhanced inputs. Furthermore, a feature selection method, Equilibrium (a one-hot encoder categorical function), is proposed to identify and prioritize the most significant features while eliminating redundancies.

This reduces computational overhead and improves classification accuracy. By addressing these challenges, the proposed framework not only enhances the efficiency and precision of mammogram image classification but also sets a foundation for more effective use of deep learning in medical image processing.

3.2. Proposed System

The research introduces a novel approach grounded in deep learning, leveraging optimal characteristics extracted from original and enhanced mammography images using the CNN-EfficientNet-B3 architecture. The study proposes a method for improving the Histogram of Oriented Gradients (HOG) based on the concept of haze removal, enhancing the clarity and feature representation of the input images. To effectively train the improved CNN-EfficientNet-B3 model, data augmentation techniques were employed, ensuring robust learning and generalization.

Hyperparameter values for the training phase were determined through systematic trials and evaluations to optimize the model's performance. The CNN-EfficientNet-B3 architecture was fine-tuned to extract deep features directly from the convolutional and average pooling layers, bypassing the traditional fully connected layer for improved feature representation. Both original and enhanced images were used for training, with the model extracting critical features from these inputs. To further refine the feature representation, a serial nature technique was employed to merge the deep features extracted during training. Additionally, a novel feature selection method, referred to as Equilibrium (a one-hot encoder categorical function), was introduced to identify and combine the most significant features.

These contributions, including the integration of the CNN-EfficientNet-B3 architecture, enhanced HOG, improved training methodology, and efficient feature selection and fusion, represent the core innovations of this work. The proposed framework demonstrates significant improvements in breast cancer detection and classification, showcasing its potential for clinical applications.

3.3. Dataset

The Mammographic Image Analysis Society (MIAS) dataset serves as a foundational resource for this study and encompasses two primary categories: cancerous (malignant) and benign tumours. Tumours are classified as benign if they lack key malignancy criteria such as significant atypia of cells, mitosis, membrane degradation, or metastasis. Each dataset entry consists of mammographic films grouped in pairs, representing the patient’s left breast (even numbers in the filenames) and right breast (odd numbers in the filenames).

Each image has dimensions of 1024x1024 pixels, with the breast region centrally aligned in the matrix for consistency. Benign tumours, referred to as "innocent," are typically slow-growing and localized, whereas malignant tumours are characterized by their ability to metastasize, damaging surrounding structures and posing severe health risks.

The MIAS dataset contains examples of both malignant and benign cases distributed across various histological subtypes, including ORC, NORM, MISC, ASYM, ARCH, SPIC, and CALC. Before analysis, preprocessing steps were undertaken to enhance the quality of the dataset.

These steps included image normalization to ensure uniform intensity distribution, resizing to ensure compatibility with the model input dimensions, and augmentation techniques to increase dataset diversity. The detailed organization and preprocessing of the MIAS dataset enabled the robust training and evaluation of the CNN-EfficientNet-B3 model, providing a reliable foundation for breast tumour detection and classification.

Table.1. Survey of deep learning based breast cancer detection methods

Reference	Scope	Technique	Dataset	No. of Sample	Training (%)	Testing (%)	Results
Parvin & Mehedi Hasan [14] (2020)	To analyze CNN models for cancer images	LeNet, VGGNet, AlexNet, InceptionV3, ResNet	BreaKHis	7909	80	20	Accuracy of 89.00%, 92.00%, 94.00% and 90.00% acquired for 40x, 100x, 200x, and 400xmagnification factors, respectively
Hameed et al. [15] (2020)	To use deep learning for the classification of breast cancer images	Variants of VGGNet (e.g., fully trained VGG16, fine-tuned VGG16, fully trained VGG19, and fine-tuned VGG19model)	Breast cancer images: 675 for training and 170 for testing	845	80	20	The ensemble offine-tuned VGG16 and VGG19 models offered a sensitivity of 97.73% for the carcinoma class and 95.29% total accuracy. It additionally provided a 95.29% F1 score.
Mahmoud et al. [16] (2021)	To analyze breast cancer images	Implemented transfer learning	Mammogram images	7500	80	20	Maximum accuracy of 97.80% was claimed by using this dataset. Sensitivity and specificity were calculated.
Boumaraf et al. [17] (2021)	To classify mammogram images	Performed ResNet on ImageNet images	BreaKHis	7909	80	20	Accuracy of 94.49%, 93.27%, 91.29%, 89.56% acquired over 40x, 100x, 200x, & 400x magnification factors, consequently
Reshma et al. [18] (2022)	To diagnose mammogram images	Utilized CNN along with probabilistic transition rules	BreaKHis	7909	90	10	Accuracy, PRS, F1S, RES, & GMN of 89.13%, 86.23%, 81.47%, 85.38%, & 85.17%
Alruwaili and Gouda [19] (2022)	To diagnose breast cancer	Implemented the transfer learning principle, ResNet	MIAS	322	80	20	Finest results for ACC, PRS, RES, F1S, and AUC of 89.5%, 89.5%, 90%, and 89.5% obtained from MIAS, respectively

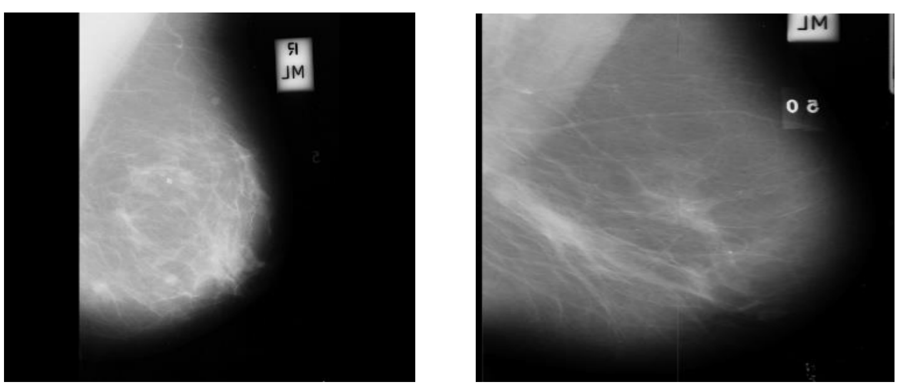

Fig. 1 Benign and malignant

3.4. Data Augmentation

For the training of Common approaches for machine learning, including form features (HOG), point features, colour features, and others, the few envision datasets are helpful. It is usually necessary to create or gather some larger datasets for deep learning models. Nevertheless, the size of the publicly accessible databases for cancer and breast images is insufficient; as a result, in this work with data augmentation. Data augmentation increases the dataset while decreasing overfitting issues and boosting the robustness of the deep learning model. Eight additional images were created for each identified patch by rotating each image four times at angles of 0°, 90°, 180°, and 270° and then flipping the four recently obtained images from right to left. Figure 2 shows the model for data augmentation for mammogram images.

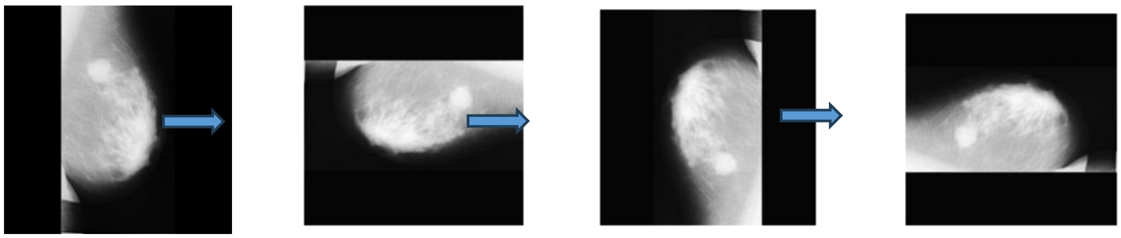

Fig. 2 Data augmentation for mammogram image

3.5. Feature Projection and Feature Scaling

Data from a space with multiple dimensions has been converted into a lower-dimensional space (with fewer properties) using feature projection. Both linear and nonlinear reduction techniques can be used, depending on how the characteristics in the dataset are correlated. The features in the data that are gathered often come in an immense range of magnitudes, ranges and units. However, given that the majority of machine learning algorithms analyze the Euclidean distance between two data points. Each attribute must be scaled up to the same magnitude. This can be attained through scaling.

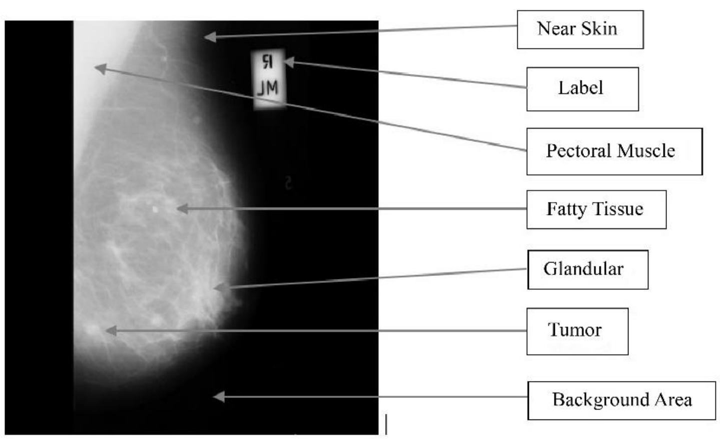

Fig. 3 Feature projection sample mammogram image

3.6. Model Selection

The supervised learning approach involves labelling the input and output of the data before the machine is trained on it. The model may be trained using historical data, and it can analyse new data to make predictions about the future. They are divided into strategies for regression and classification. When the outcome is a continuous or real value, like "weight" or "salary," a regression challenge exists. When the outcome is a category like screening emails as "not spam" or "spam," there is a classification difficulty. Unsupervised learning is the process of giving the computer data that has not been categorised or marked and allowing the algorithm to study the data without being provided any instructions. The computer is taught from the data in an unsupervised learning technique. It is missing labels or classification, causing the algorithm to operate without sufficient direction. The outcome variable, or dependent variable, Y, in our dataset only has two possible values: M (Malign) or B (Benign). Therefore, the supervised learning Classification method is used on it. Three main categories of machine learning classification algorithms have been selected. The work can employ the simple little linear model.

3.7. CNN-Efficientnet B3 Algorithm

The proposed model consists of a CNN and transfer learning based on EfficientNetB3 (CNN-EfficientNetB3). The input image for this model is 1024 x 1024 pixels, with 5,07,21,329 parameters. Table 2 illustrates the suggested model's overview. CNN for our application from scratch.

A CNN's network layers serve as a detection filter for particular patterns or features that may be present in an image. Large features that are simple to understand are found in the first levels of a CNN. The lower levels find the more abstract, smaller features. By integrating all the features identified by preceding layers, the last layer is able to generate a reasonably thorough classification.

Fig. 4 CNN architecture diagram

The customized CNN model architecture consists of multiple convolutional layers, activation functions, batch normalization, and max pooling operations. Convolutional layers make up the first four of CNN's seven weighted layers, while fully connected layers make up the final three. "Grey scale images are the DCNN's inputs. To the local area that is related to the input quantity, each neuron computes a weighted dot product. This work utilised padding of sizes (3, 2, 1) along the outermost edges of the input layer and 4, 16, and 80 filters of sizes (2, 3, 5). The filters defined by the filter size (3, 3) have a height and length of 3. The width and height of each filter are moved across the input. Down sampling is performed by two pooling layers, which helps to reduce computation and improve robustness. Pooling layers with a 2 m filter by 2 pixels, "outputting" the highest value achievable from each local region's 4 inputs. A CNN-based classifier's final layer is usually the Softmax Layer.

Table 2. Summary of the proposed model

Layers	K_Size	Input	Act_Function	Output
Conv2D_1 Max_pooling_1 Dropout=0.5	3X3 2X2 -----	(32,32,32) (32,32,32) (16,16,32)	Relu ----- -----	(32,32,32) (16,16,32) (16,16,32)
Conv2D_2 Dropout=0.5	3X3 -----	(16,16,32) (16,16,64)	Relu -----	(16,16,64) (16,16,64)
Conv2D_3 Max_pooling_3 Dropout=0.5	3X3 2X2 -----	(16,16,64) (16,16,64) (8,8,64)	Relu ----- -----	(16,16,64) (8,8,64) (8,8,64)
Conv2D_4 Dropout=0.5	3X3 -----	(8,8,64) (8,8,128)	Relu -----	(8,8,128) (8,8,128)
Conv2D_5 Dropout=0.5	3X3 -----	(8,8,128) (8,8,128)	Relu -----	(8,8,128) (4,4,128)
Conv2D_6 Max_pooling_6 Dropout=0.5	3X3 2X2 -----	(4,4,128) (4,4,256) (2,2,256)	Relu ----- -----	(4,4,256) (2,2,256) (2,2,256)
Flatten_CNN Flatten_1_eff	----- -----	(2,2,256) (1,1,1024)	----- -----	1024 1280
Dense1 Dense2 Output	----- -----	1280 64 16	Sigmoid Sigmoid	64 16 2

Each epoch's weight changes are impacted by the learning rate; for example, a higher learning rate leads to greater weight changes, which speed up network learning and vice versa. A learning rate of 0.01 has been employed. EfficientNetB3 uses a compound coefficient to reduce the convolutional neural network's size and architecture. This compound maintains stability across the network. The coefficient adjusts the depth, breadth, and resolution parameters regularly. The EfficientNetB3 scaling methodology, in contrast to the traditional approach, uses a predetermined collection of scaling coefficients that constantly change the depth, breadth and resolution of the network. According to the complex scaling method, this network required extra layers to increase its channels and receptive field to identify fine aspects inside the larger image as input images get higher.

Fine-Tuned Model: The EfficientNet-B3 network was initially trained and utilizing the ImageNet dataset, which consists of images from more than a million distinct object categories. The completely linked, softmax, and output layers were where the first three new levels were introduced; the final three layers were left out for fine-tuning.

Subsequently, the stochastic gradient descent optimizer is set up with a momentum of 0.703, a learning rate of 0.005, and a total of 100 epochs. Finally, this improved model uses the deep learning transfer Model. Transfer learning is the process by which knowledge is transferred from one subject or domain to another. Especially in the area of medical imaging, it is challenging and time-consuming to learn precise patterns using deep learning. Employed a specific source as the "source" and used a labelled dataset.

3.8. Fine-Tuning Parameters and Overfitting Mitigation Strategies

To ensure optimal performance and reproducibility of the CNN-EfficientNet-B3 model for mammogram image classification, specific fine-tuning parameters and strategies to address overfitting were implemented. A learning rate of 0.001 was selected to balance efficient convergence with stability, further refined using a learning rate scheduler to adaptively reduce the rate as the model approached optimal weights.

A batch size of 32 was utilized, ensuring efficient memory usage and gradient stability. To enhance dataset diversity and improve model generalization, various data augmentation techniques were applied, including random rotations up to 15 degrees, horizontal and vertical flips, random cropping, intensity adjustments for brightness and contrast, and the addition of Gaussian noise.

These augmentations enabled the model to learn robust features from diverse representations of mammography images. To mitigate overfitting, several regularization techniques were integrated into the model architecture and training process. Dropout layers with a rate of 0.3 were added to the fully connected layers to prevent neuron co-adaptation by randomly disabling a fraction of them during training. Early stopping was employed, monitoring validation loss and halting training after 10 epochs of no improvement to avoid overfitting the training data. Additionally, L2 regularization (weight decay) with a factor of 0.0001 was applied to penalize large weight magnitudes, promoting smoother learning. Batch normalization layers were incorporated throughout the network to standardize intermediate feature maps, reduce internal covariate shifts, and improve training stability. These carefully selected fine-tuning parameters and robust regularization strategies ensured a balance between high accuracy and generalization, making the framework both effective and reproducible for mammogram image classification.

3.9. Potential Limitations of the MIAS Dataset

The MIAS dataset, while valuable for breast cancer research, has several limitations that could affect the generalizability and robustness of models trained on it. One of the primary concerns is its relatively small size, which limits the model’s ability to generalize to larger, more diverse datasets. Additionally, the lack of demographic diversity in the dataset introduces potential biases, as it predominantly represents a specific population, reducing the model’s effectiveness when applied to different populations with varying breast tissue densities, imaging techniques, or genetic backgrounds.

There is also the issue of class imbalance, as the distribution of malignant and benign cases may not be equally represented, leading to biased performance towards the majority class and reduced sensitivity in detecting the minority class. Furthermore, the dataset lacks accompanying clinical metadata, such as patient age, medical history, or breast density, which could provide valuable context for improving model interpretability and accuracy. The dataset’s low-resolution images (1024x1024 pixels) also limit the detection of fine details that may be crucial for accurate diagnosis, especially when compared to modern high-resolution mammography systems.

Additionally, the pairing of left and right breast images could introduce correlation bias, potentially influencing the model’s learning if one side contains a malignant tumour. Lastly, the absence of detailed annotations for subtle features like microcalcifications or architectural distortions makes it difficult for the model to learn these critical indicators of early-stage cancer.

These limitations highlight the need for careful consideration during model development, and strategies such as data augmentation, transfer learning, and external validation on larger, more diverse datasets can help mitigate these constraints and enhance model generalizability.

4. Results and Discussion

4.1. Experimental Setup

This model was executed using Tensorflow with the Keras package in Python. The experiment's configuration used 12 GB RAM and a Tesla K80 GPU, disk space of about 78 GB, which was run on Google Colab. The obtained data was 1024 by 1024 in size.

First, split the dataset into two classes, benign and malignant, then used 150 and 100 of the generated images for training, while the leftover images were used for testing. Different size filters (sizes 2, 3, and 5) were used.

Additionally, the work attempted to manually and automatically divide each training and testing dataset by a 70:30 ratio, but the outcomes varied. When data is trained randomly, it produces better outcomes than when it is given instruction automatically.

4.2. Performance Metrices

The F1 score is a metric used in statistical analyses of binary classification to assess the accuracy of a test. It is also known as an F-score or F-measure. The total number of appropriate samples (all examples that should have been categorized as positive) and the total number of positive outcomes returned by the classifier are divided to figure out the score. The proposed methodology performance is examined with the help of the following metrics:

Accuracy, Area Under the Curve (AUC), Receiving Operating features (ROC), Precision, and F1 score features. According to the previously suggested relationship, classifier accuracy.

Classifier Accuracy= TP+TN/TN+TP+FP+FN

Mapping the False Positive Rate (FPR) at the X-axis as well as the True Positive Rate (TPR) at the Y-axis results in a Receiver Operating Characteristic (ROC) graph. The calculated specificity and sensitivity of the classifier are frequently referred to as the false positive rate and the true positive rate.

Sensitivity= TP/TN+FN

Specificity= TN/TN+FP

TP and FP rates are determined utilizing the Area Under the Curve (AUC) technique under the ROC curve. One can calculate the accuracy of actual positive predictions.

Prediction= TP/TP+FP

The weighted average of sensitivity and precision, known as the F1 Score, is applied to determine classifier performance.

$$F1\ Score = \frac{2 * Sensitivity * Precision}{Sensitivity + Precision}$$

```
[ ] sns.set_style('darkgrid')
    fig, (ax1, ax2) = plt.subplots(1,2,figsize=(15,5))
    sns.barplot(x=info.BG.unique(),y=info.BG.value_counts(),palette='pink',ax=ax1)
    sns.barplot(x=info.CLASS.unique(),y=info.CLASS.value_counts(),palette='pink',ax=ax2)
```

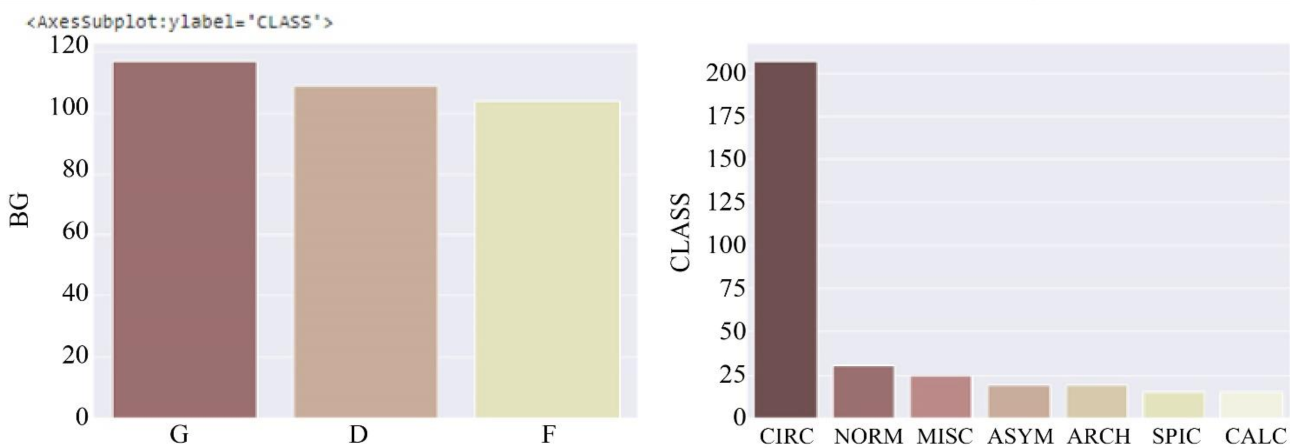

Fig. 5 Historical representation

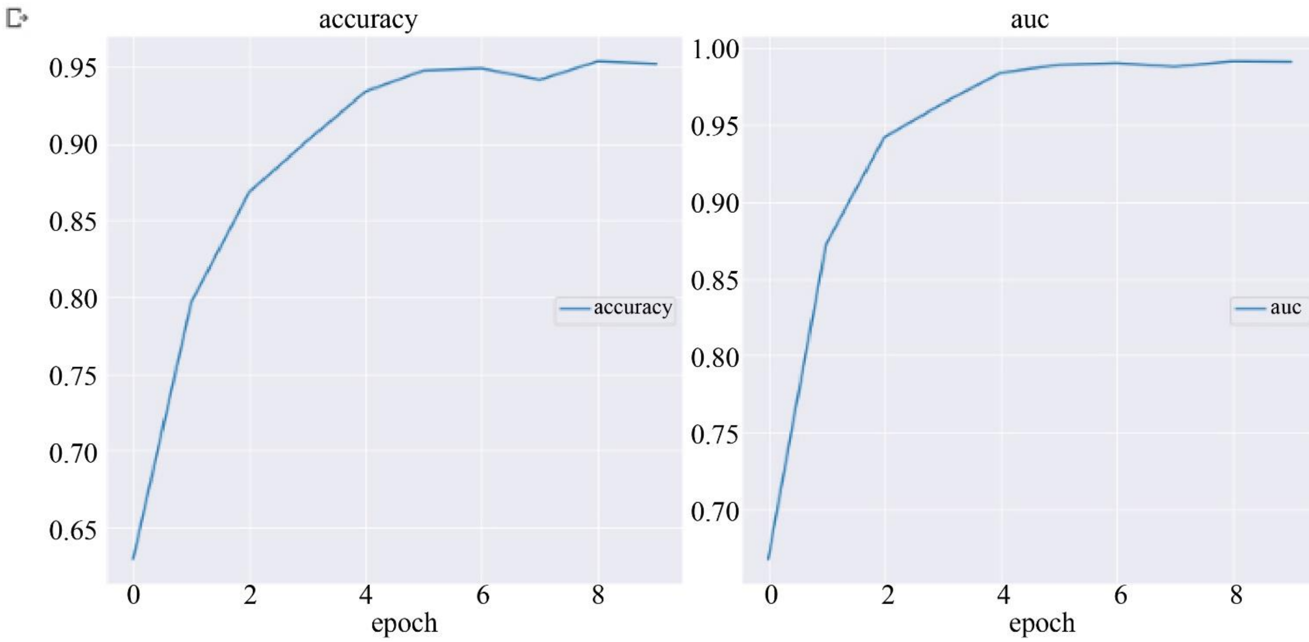

Fig. 6 Accuracy and AUC curve for CNN-EfficientNetB3 model

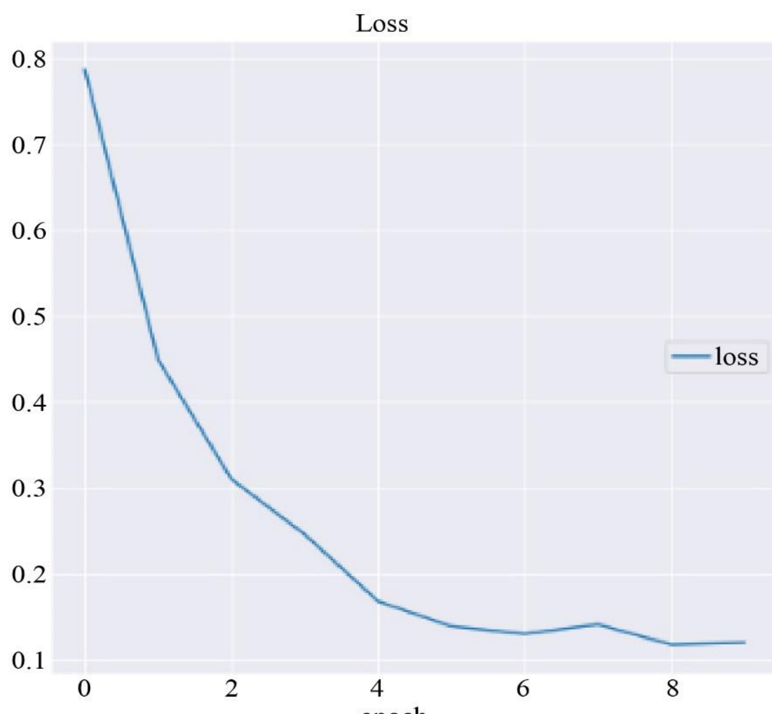

Fig. 7 Loss curve for CNN-EfficientNetB3 model

4.3. Performance Metrics

The experimental study was conducted by the CNN-EfficientNetB3. Figure 5 represents the historical representation of the histological dataset with its count and class. Figure 6 shows the accuracy AUC curve for the CNN-EfficientNetB3 model, and Figure 7 shows the loss curve for the same model. Figure 8 shows the performance analysis of this model during prediction. The outcomes demonstrate the accuracy of the CNN + EfficientNetB3 model, which is 97% accuracy,97.50% recall, and also 97% F1 score. According to tests, the CNN + EfficientNetB3 model has the highest accuracy compared with the CNN model and theother deep learning models that are shown in Table 3.

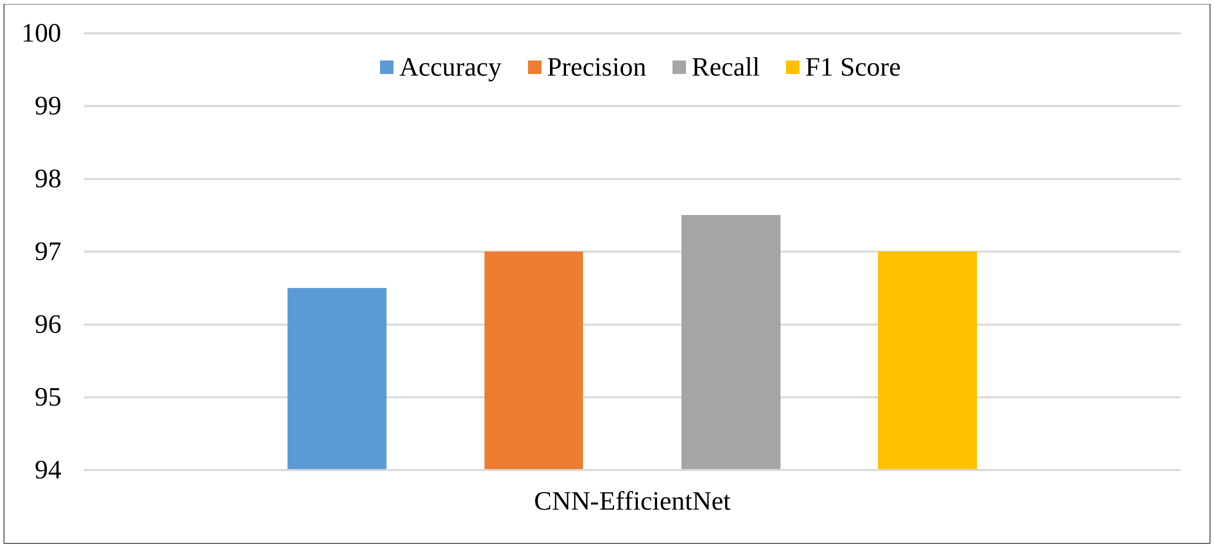

Fig. 8 Performance analysis for the proposed method

```
WARNING:tensorflow:Learning rate reduction is conditioned on metric `val_loss` which is not available. Available metrics are: loss,accuracy,auc,lr
accuracy
        accuracy             (min:    0.629, max:    0.954, cur:    0.952)
auc
        auc                  (min:    0.668, max:    0.991, cur:    0.991)
Loss
        loss                 (min:    0.118, max:    0.786, cur:    0.120)
227/227 [==============================] - 9s 38ms/step - loss: 0.1201 - accuracy: 0.9522 - auc: 0.9909 - lr: 0.0010

model.evaluate(x_val,y_val,callbacks=[c2,c3],batch_size=16)

68/68 [==============================] - 5s 37ms/step - loss: 0.1015 - accuracy: 0.9623 - auc: 0.9931
[0.10147206485271454, 0.9622815251350403, 0.9931336641311646]
```

Fig. 9 Accuracy prediction

Table 3. Performance comparison between the proposed model with earlier models

Author	Approach	Accuracy
Gupta et al. [20]	CNN	87%
Barsha et al. [21]	DenseNet-121 and DenseNet-169	92.70%
Humayun et al. [22]	InceptionResNetV2	91%
Abdolahi et al. [23]	CNN	85%
Zhang et al. [24]	Alexnet, MobilenetV2, and Resnet50	87.45%
Wang et al. [25]	CNN-GRU	86.21%
Celik et al. [26]	ResNet-50 and DenseNet-161	91.57%
Singh et al. [27]	DenseNet + Logistic Regression	81%
Kundale et al. [28]	GoogLeNet+ CNN	94%
Proposed Model	CNN - EfficientNetB3	96.5%

In this study, several evaluation metrics were used to assess the performance of the CNN-EfficientNet-B3 model for breast cancer classification, including accuracy, sensitivity, and specificity. However, to provide a more comprehensive evaluation, additional metrics such as F1 score, precision, and ROC-AUC were also considered. Accuracy measures the overall proportion of correct predictions, but it may not be fully reliable in cases of class imbalance. Sensitivity (or recall) evaluates the model's ability to correctly identify true positive cases of malignant tumours, which is critical in medical diagnostics to avoid false negatives. Specificity, on the other hand, assesses the model's ability to correctly identify true negatives (benign cases), ensuring that benign tumours are not misclassified as malignant.

Precision measures the proportion of true positives among all positive predictions, which is important to minimize false positives and prevent unnecessary treatments. The F1 score, being the harmonic mean of precision and recall, offers a balanced evaluation and is particularly useful in imbalanced datasets, ensuring the model maintains a good balance between identifying true positives and avoiding false positives. Lastly, ROC-AUC provides a graphical representation of the model's ability to distinguish between malignant and benign cases across various classification thresholds, with a higher AUC indicating better overall performance. By incorporating these additional metrics, a more holistic evaluation of the model's performance is achieved, providing deeper insights into both its strengths and areas for improvement, especially in medical applications where the consequences of false positives and false negatives are significant.

5. Conclusion

This study proposes an entirely novel approach for classifying or predicting breast cancer using mammography images. The suggested structure starts with imagine acquisition and classification of images and comprises essential components. In this session, a method for contrast enhancement is recommended. The upgraded images were needed to train the deep learning model (CNN-EfficientNetB3), and the outcomes were compared to the precision of the deep features in the original image.

Results indicate the accuracy of our proposed or enhanced model is higher than the accuracy recently achieved; however, as a result, a new Combine technique has been Created. The raw image and augmented image features were combined using the recommended fusion model, which produced an amazing improvement in accuracy.

Author Contribution

Punithavathi K conceptualized the study, developed the methodology, and supervised the work. She led the design and implementation of the machine learning models and played a significant role in writing the manuscript. G. Yamuna was responsible for data analysis, the design and deployment of the high-definition microphone system, and the development of algorithms. She also assisted in manuscript preparation and editing, contributed to the literature review and background research, coordinated research activities across different institutions, and supported manuscript drafting and proofreading.

References

[1] Ihsan Jasim Hussein et al., "Fully-Automatic Identification of Gynaecological Abnormality Using a New Adaptive Frequency Filter and Histogram of Oriented Gradients (HOG)," *Expert Systems*, vol. 33, no. 3, pp. 1-47, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[2] Seifedine Kadry et al., "Extraction of Tumour in Breast MRI using Joint Thresholding and Segmentation-A Study, *2021 Seventh International Conference on Bio Signals, Images, and Instrumentation (ICBSII)*, Chennai, India, pp. 1-5, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[3] Rizwana Irfan et al., "Dilated Semantic Segmentation for Breast Ultrasonic Lesion Detection Using Parallel Feature Fusion," *Diagnostics*, vol. 11, no. 7, pp. 1-20, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[4] Venkatesan Rajinikanth et al., "Breast Cancer Detection Using Thermal Images with Marine-Predators-Algorithm Selected Features," *2021 Seventh International Conference on Bio Signals, Images, and Instrumentation (ICBSII)*, Chennai, India, pp. 1-6, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[5] Ihsan Jasim Hussein et al., "Fully Automatic Segmentation of Gynaecological Abnormality Using a New Viola-Jones Model," *Computers, Materials & Continua*, vol. 66, no. 3, pp. 3161-318, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[6] Katharina Holland et al., "Quantification of Masking Risk in Screening Mammography with Volumetric Breast Density Maps," *Breast Cancer Research and Treatment*, vol. 162, pp. 541-548, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[7] Adrian P. Brady, "Error and Discrepancy in Radiology: Inevitable or Avoidable?," *Insights into Imaging*, vol. 8, pp. 171-182, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[8] Piero Olliaro, and Els Torreele, "Managing the Risks of Making the Wrong Diagnosis: First, do No Harm," *International Journal of Infectious Diseases*, vol. 106, pp. 382-385, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[9] Saleem Z. Ramadan, "Methods Used in Computer-Aided Diagnosis for Breast Cancer Detection Using Mammograms: A Review," *Journal of Healthcare Engineering*, vol. 2020, no. 1, pp. 1-21, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[10] Erich Sorantin, "The Augmented Radiologist: Artificial Intelligence in the Practice of Radiology," *Pediatric Radiology*, vol. 52, pp. 2074-2086, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[11] Gelan Ayana, Kokeb Dese, and Se-woon Choe, "Transfer Learning in Breast Cancer Diagnoses via Ultrasound Imaging," *Cancers*, vol. 13, no. 4, pp. 1-15, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[12] Hanan Aljuaid et al., "Computer-Aided Diagnosis for Breast Cancer Classification Using Deep Neural Networks and Transfer Learning," *Computer Methods and Programs in Biomedicine*, vol. 223, pp. 1-10, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Y. Wang et al., "Improved Breast Cancer Histological Grading Using Deep Learning," *Annals of Oncology*, vol. 33, no. 1, pp. 89-98, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[14] Farjana Parvin, and Md. Al Mehedi Hasan, "A Comparative Study of Different Types of Convolutional Neural Networks for Breast Cancer Histopathological Image Classification," *2020 IEEE Region 10 Symposium (TENSYMP)*, Dhaka, Bangladesh, pp. 945-948, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[15] Zabit Hameed et al., "Breast Cancer Histopathology Image Classification Using an Ensemble of Deep Learning Models," *Sensors*, vol. 20, no. 16, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[16] Hanan A. Hosni Mahmoud, Amal H. Alharbi, and Doaa S. Khafga, "Breast Cancer Classification Using Deep Convolution Neural Network with Transfer Learning," Intelligent Automation & Soft Computing, vol. 29, no. 3, pp. 803-814, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[17] Said Boumaraf et al., "A New Transfer Learning Based Approach to Magnification Dependent and Independent Classification of Breast Cancer in Histopathological Images," *Biomedical Signal Processing and Control*, vol. 63, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[18] V.K. Reshma et al., "Detection of Breast Cancer Using Histopathological Image Classification Dataset with Deep Learning Techniques," *BioMed Research International*, vol. 2022, no. 1, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Madallah Alruwaili, and Walaa Gouda, "Automated Breast Cancer Detection Models Based on Transfer Learning," *Sensors*, vol. 22, no. 3, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[20] Isha Gupta, Sheifali Gupta, and Swati Singh "Architectures Based on Deep Learning for the Detection of Invasive Ductal Carcinoma," *ECS Transactions*, vol. 107, no. 1, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[21] Nusrat Ameen Barsha, Aimon Rahman, and M.R.C. Mahdy, "Automated Detection and Grading of Invasive Ductal Carcinoma Breast Cancer Using Ensemble of Deep Learning Models," *Computers in Biology and Medicine*, vol. 139, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[22] Mamoona Humayun et al., "Framework for Detecting Breast Cancer Risk Presence Using Deep Learning," *Electronics*, Vol. 12, no. 2, pp. 1-16, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[23] Mohammad Abdolah et al., "Artificial Intelligence in Automatic Classification of Invasive Ductal Carcinoma Breast Cancer in Digital Pathology Images," *Medical Journal of the Islamic Republic of Iran*, vol. 34, no. 140, pp. 1-9, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[24] Jianfei Zhang et al., "Automatic Detection of Invasive Ductal Carcinoma Based on the Fusion of Multi-Scale Residual Convolutional Neural Network and SVM," *IEEE Access*. vol. 9, pp. 40308-40317, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[25] Xiaomei Wang et al., "Intelligent Hybrid Deep Learning Model for Breast Cancer Detection," *Electronics*, vol. 11, no. 17, pp. 1-18, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[26] Yusuf Celik et al., "Automated Invasive Ductal Carcinoma Detection Based Using Deep Transfer Learning with Whole-Slide Images," *Pattern Recognition Letters*, vol. 133, pp. 232-239, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[27] Sukhendra Singh et al., "Hybrid Models for Breast Cancer Detection via Transfer Learning Technique," *Tech Science Press*, vol. 74, no. 2, pp. 3063-3083, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[28] Jyoti Kundale, and Sudhir Dhage, "Classification of Breast Cancer Using Histology Images: Handcrafted and Pre-Trained Features-Based Approach," *IOP Conference Series: Materials Science and Engineering*, vol. 1074, pp. 1-18, 2021. [CrossRef] [Google Scholar] [Publisher Link]