Research Article GU J Sci, Part A, 11(4): 647-667 (2024) 10.54287/gujsa.1529857

Gazi University
Journal of Science
PART A: ENGINEERING AND INNOVATION
http://dergipark.org.tr/gujsa

Advanced CNN-Based Classification and Segmentation for Enhanced Breast Cancer Ultrasound Imaging

Jehad CHEYI[1*] Yasemin ÇETIN-KAYA[1]

[1] *Tokat Gaziosmanpaşa University, Department of Computer Engineering, Tokat, Türkiye*

Keywords

Deep Learning
CNN
Breast Cancer
Classification
Image Processing

Abstract

Breast cancer (BC) is one of the primary causes of mortality in women globally. Thus, early and exact identification is critical for effective treatment. This work investigates deep learning, more especially convolutional neural networks (CNNs), to classify BC from ultrasound images. We worked with a collection of breast ultrasound images from 600 patients. Our approach included extensive image preprocessing techniques, such as enhancement and overlay methods, before training various deep learning models with particular reference to VGG16, VGG19, ResNet50, DenseNet121, EfficientNetB0, and custom CNNs. Our proposed model achieved a remarkable classification accuracy of 97%, significantly outperforming established models like EfficientNetB0, MobileNet, and Inceptionv3. This research demonstrates the ability of advanced CNNs, when paired with good preprocessing, to significantly enhance BC classification from ultrasound images. We further used Grad-CAM to make the model interpretable so we may see which parts of the images the CNNs focus on when making decisions.

Cite

Cheyi, J., & Çetin-Kaya, Y. (2024). Advanced CNN-Based Classification and Segmentation for Enhanced Breast Cancer Ultrasound Imaging. *GU J Sci, Part A, 11*(4), 647-667. doi:10.54287/gujsa.1529857

Author ID (ORCID Number)

0009-0003-5407-569X Jehad CHEYI
0000-0002-6745-7705 Yasemin ÇETIN-KAYA

Article Process

Submission Date 08.08.2024
Revision Date 16.09.2024
Accepted Date 07.10.2024
Published Date 31.12.2024

1. INTRODUCTION

The most prevalent cancer among women globally is BC (T. Wu et al., 2019). For women, it ranks second most-often occurring cause of death (Luo et al., 2024). BC affects roughly 2.89 million women globally each year, including 24.2% of all female cancer cases and placing first (Kabir et al., 2021; Pacal, 2022). Early detection of the disease is essential since it aids in lowering the number of premature deaths (G. G. Wu et al., 2019; Pourasad et al., 2021; Jabeen et al., 2022; Cruz-Ramos et al., 2023). Over 30 percent of all cancer fatalities are caused by BC, making it the deadliest cancer for women (Vigil et al., 2022). BC is identified with a variety of imaging methods, including mammography, ultrasound, scanning with magnetic resonance imaging (MRI), and electronic pathology images. Mammography is widely utilized for early detection. However, it has limits in sensitivity and specificity, particularly in thick breasts (Badawy et al., 2021). For breast ultrasound images, state-of-the-art CNN techniques like Single Shot Multibox Detector with YOLO, You Only Look Once have shown significant success in detecting breast lesions (Fujioka et al., 2020). MRI provides detailed images of the breast and is especially helpful in assessing high-risk patients and assessing tumor extent (Peng et al., 2023). Digital pathology pictures are the "golden standard" for recognizing cancer and are critical in cancer detection. Deep learning has demonstrated encouraging outcomes in increasing BC detection and categorization precision and efficiency. Deep learning has also been utilized to identify BC via ultrasound images, where data augmentation and transfer learning find application approaches have been applied to increase performance (G. G. W. et al., 2019). CNNs and deep learning techniques generally, have been used to assess medical images, which are mammography, ultrasound, MRI, and images from pathology

*Corresponding Author, e-mail: jehad_cheyi@yahoo.com

(Mahoro & Akhloufi, 2022). Deep learning has significantly improved diagnosis accuracy and treatment efficiency in a variety of medical areas (Çetin-Kaya & Kaya, 2024; Kaya & Cetin-Kaya, 2024). Deep learning algorithms can extract meaningful characteristics from imagery and make classifications based on them (Luo et al., 2024). For example, in mammogram-based screening and diagnosis, deep learning models have been used to classify mammograms into different categories, such as malignant, benign, or normal (Y. Zhang et al., 2021). Transfer learning involves pre-training a structure on a big dataset and adjusting it for a specific job. It has also been employed to enhance the efficiency of deep learning algorithms in BC imaging (Luo et al., 2024). There are challenges, such as the need for interdisciplinary cooperation, standardized and open databases, and addressing deep learning models' poor generalization ability and interpretability (Han et al., 2017). In this article, we explore the transformative potential of deep learning in BC ultrasound imaging, characterized by a diversity of innovative solutions that extend far beyond classification, with the development of a novel strategy and solutions. Current imaging techniques, such as mammography, often fall short in sensitivity and specificity, especially in women with thick breast tissue. This study seeks to overcome these constraints by investigating sophisticated approaches such as deep learning, notably CNNs, to improve diagnostic precision and help radiologists in ultrasound imaging. Our research seeks to enhance BC diagnosis utilizing ultrasound pictures and powerful deep-learning algorithms. Our contributions are as follows:

- Developed a custom CNN model achieving 97% classification accuracy for BC detection in ultrasound images, outperforming various state-of-the-art algorithms.
- Applied advanced image processing techniques, including U-Net for segmentation and Grad-CAM for model interpretability, to improve diagnostic accuracy and transparency.

The following describes the latter bits of this research. Section 2 addresses related works; Section 3 offers the dataset and approach of the investigation. Section 4 offers the findings of the research. Section 5 presents the findings of the study together with suggestions for next investigations.

2. RELATED WORK

The related work analysis in breast ultrasonography covers a broad spectrum of deep learning and transfer learning applications for classification, diagnosis, detection, and segmentation. Researchers have proposed novel deep neural network architectures and inventive approaches to enhance automated breast ultrasound systems' performance, specificity, and accuracy. Their combined efforts improve the possibility of an early diagnosis and better patient outcomes by advancing the detection and evaluation of BC. Table 1 summarizes previous research on the use of ultrasonography to diagnose BC.

2.1. Classification

In breast ultrasound classification, researchers have used deep-learning models to achieve remarkable accuracy, frequently outperforming traditional methods. These models demonstrate their ability to identify normal, benign, and malignant cases, providing invaluable assistance to radiologists and improving diagnostic accuracy. Classification-focused research contributes significantly to improving automated BC diagnosis and risk assessment.

2.1.1. Deep Learning

Several studies on BC have made significant contributions to diagnosis and classification by implementing deep learning techniques. Jabeen et al. (2022) introduced a probability-based optimum deep learning feature fusion method, achieving an impressive accuracy of 99.1% in BC classification. Their method utilized the Breast Ultrasound Images (BUSI) dataset, including 780 images categorized as normal, malignant, or benign. Zhuang et al. (2021) took a different approach, using image decomposition and fusion, including adaptive multi-model spatial feature fusion. This method yielded a remarkable accuracy of 95.48% and high precision on BUSI dataset. Momot et al. (2022) utilized deep neural networks, the EfficientNet B0, pre-trained using the data set from ImageNet, to automate ultrasound BC image classification. They reached an accuracy of 81.26%, highlighting the promise of deep learning in ultrasound diagnostics. Alrubaie et al. (2023) implemented CNN and transfer learning to achieve a high classification accuracy, with a 96% accuracy rate for one dataset and a

perfect 100% accuracy for another. This study involved datasets divided into Group A, comprising 780 images with three classes (benign, malignant, normal), and Group B, consisting of 9,016 images with benign and malignant BCs. A noise filter network (NF-Net) was used by Cao et al. (2020) to learn from noisy tagged ultrasound images to classify breast tumors. This approach attained classification accuracy of 73%. Furthermore, Kim et al. (2021) provided research regarding deep learning with little supervision for detecting ultrasounds of BC. Moon et al. (2020) focus on the diagnosis of BC. The researchers achieved high accuracy and AUC values using convolutional neural networks for ensemble learning. They used a private dataset containing 1687 tumors, with 953 benign and 734 malignant cases. Along with an open BUSI dataset with 697 tumors, comprising 437 benign, 210 malignant, and 133 normal cases. The study presents a computer-aided diagnostic (CAD) system that utilizes CNN architectures for tumor diagnosis. Liu et al. (2022) performed BC diagnosis using artificial intelligence and deep learning techniques. The suggested grid-based deep characteristic generator model classified breast ultrasonic images with 97.18% accuracy across cancerous, benign, and normal classes. The performance measures include accuracy, recall, precision, F1-score, and geometric mean, exceeded 96 for all classes. The study used a BUSI dataset comprising images from 600 female patients. The method involves grid-based deep feature generation, pre-developed CNN models, incremental feature selection, and a deep classifier to improve BC diagnosis accuracy.

2.1.2. Transfer Learning for Breast Cancer Diagnosis

G. G. Wu et al. (2019) focus on using machine learning, particularly transfer learning, for diagnosing triple-negative BC with ultrasonography. The dataset consists of 140 surgically confirmed BC patients, incorporating both ultrasound and clinical data. Grayscale and color doppler features were utilized for classification. The research demonstrated that machine learning using quantified ultrasonic image characteristics, including color doppler information, effectively differentiates triple-negative BC cases. Moustafa et al. (2020) used color doppler ultrasound introduced to enhance BC detection using machine learning. The dataset contains 159 solid masses, with 95 benign and 64 malignant cases. The study illustrates that incorporating color doppler and grayscale features in the training dataset improves the receiver operating characteristic area, thus enhancing the BC prognosis. Qi et al. (2019) proposed an automated diagnosis method for deep neural networks to process breast ultrasound pictures. Their approach employs a cascade of two neural networks, Mt-Net and Sn-Net, for breast ultrasonography image diagnosis. Mt-Net classifies images for the presence of malignant tumors, while Sn-Net further classifies images for solid nodules. The study used a large-scale ultrasound image collection annotated by experts separated into training, validation, and testing sets. The proposed method demonstrated comparable performance to human sonographers and achieved high accuracy and specificity in BC diagnosis. Several studies have harnessed transfer learning techniques to improve BC diagnosis and classification in the context of ultrasound imaging. Y. Zhang et al. (2021) investigated the use of transfer learning to fine-tune a model for improved BC subtype prediction. The study showed that transfer learning significantly enhanced classification accuracy.

Similarly, Pang et al. (2021) augmented breast ultrasound mass classification data using a semi-supervised Generative Adversarial Network(GAN)-based radiomics model. This approach resulted in an accuracy of 90.41% and high sensitivity, demonstrating the potential of transfer learning in data augmentation. Coronado-Gutiérrez et al. (2019) studied quantification ultrasound evaluation of images to diagnose Metastatic BC invasion. The suggested approach obtained an accuracy of 86.4% and a sensitivity of 84.9% using a dataset of 105 patients submitted 118 lymphatic node ultrasound images selected from two hospitals. Cruz-Ramos et al. (2023) applied transfer learning on a pre-trained architecture to classify benign and malignant BCs in ultrasound and mammography pictures, resulting in 97.6% accuracy. These studies' results show that transfer learning methods improve the efficiency of BC diagnosing and classification techniques using ultrasound imaging.

2.2. Detection using Deep Learning

Several research has investigated the utilization of Deep learning algorithms for detecting BC in radiography and ultrasound imaging. Mahoro and Akhloufi (2022) evaluated therapy response using reference images from

the Evaluating Therapy Response Breast MRI dataset (1,500 DICOM format images) and the BUSI dataset (250 BMP type BC images). Marini et al. (2023) introduced a method called volume sweep imaging for BC detection. Masud et al. (2021) uses CNNs that have been previously trained to identify BC using ultrasound imaging. Their study demonstrated impressive results, with models like DenseNet201, Xception, and ResNet18 achieving 100% accuracy using various optimizers. This study used publically accessible breast ultrasound scans from Rodrigues; there are 250 ultrasound scans, including 100 normal and 150 cancerous instances. The study applied a transfer learning technique to pre-trained models and a K-fold cross-validation procedure for assessment. Li et al. (2022) presented the "BUSnet" Deep learning algorithm for breast tumors lesion detection in ultrasound images. Their dataset included 780 samples, with 133 normal, 487 benign, and 210 malignant cases. The testing dataset was collected from a prior source. The method included unsupervised region proposal, bounding-box regression methods, and a post-processing strategy to improve detection accuracy. Z. Zhang et al. (2021) focused on employing a Deep learning algorithm used in automated breast ultrasonography to detect cancers. They achieved a sensitivity of 0.88 with a false positive rate of 0.19 per second (FP/S). The study employed a unique ABUS imaging dataset from Peking University People's Hospital, which included 170 ABUS tumor volumes obtained from 124 female patients. Their technique comprises employing the Bayesian YOLOv4 network and Monte Carlo dropout and making unique YOLOv4 tweaks for detecting ABUS tumors.

2.3 Segmentation

2.3.1. Deep Learning-Based Segmentation Studies

Gómez-Flores and de Albuquerque Pereira (2020) focus on evaluating the performance of pre-trained CNNs at segmenting BCs in ultrasound pictures. They revealed that SegNet and DeepLabV3+ had the greatest segmentation results, while ResNet18 showed potential for CAD systems. The study included more than 3000 breast ultrasound Images taken from seven separate ultrasound machine models. Ilesanmi et al. (2021) suggested a segmentation strategy producing high dice measures for malignant and normal breast ultrasound image. The study used two datasets: one with 264 photos (100 malignant, 164 benign) and another with 830 images (487 malignant, 210 benign, and 133 normal). Their approach includes a preprocessing stage and a deep learning segmentation stage. Vakanski et al. (2020) incorporating attention blocks into deep-learning models for breast tumor segmentation resulted in models that outperformed the basic U-Net model. The dataset included 510 ultrasound images converted to grayscale 8-bit data and resampled into floating points using normalization. Xu et al. (2019) primarily focused on applying machine learning to segment breast ultrasound pictures, with quantitative measurements achieving more than 80% accuracy. The dataset utilized is not specified. However, their method categorizes BUSI into four tissue types: skin, fibroglandular, mass, and fatty tissue. Zhang et al. (2019) introduced a novel method, Boundary-aware Semi-Supervised Deep Learning (BASDL), for breast ultrasound lesion CAD. BASDL achieved a classification accuracy of around 92.00±2.38% and was evaluated on two breast ultrasound datasets. Chiang et al. (2018) specifically targeted tumor detection in automated ultrasound image with a dataset that includes 230 pathology-proven lesions from 187 individuals, 90 of which are benign and 140 of which are malignant. The proposed method involved preprocessing, segmentation, and feature extraction using a sliding window detector for localized analysis. Lei et al. (2018) obtained the highest segmentation performance by developing a technique for segmenting breast anatomy in full breast ultrasound images. The ConvEDNet is a deep convolutional encoder-decoder network with regularized boundaries. The limited dataset used manual annotation due to the associated cost. Gong et al. (2020) explore a bi-modal approach to BC diagnosis using A support vector machine with several views based on deep neural networks. The classification accuracy achieved was 86.36%, with an AUC of 0.9079. The dataset utilized in this study included a total of 264 pairings of breast ultrasound and ultrasound elastography images of 129 individuals with benign tumors and 135 people with malignant tumors. The suggested model is divided into two parts: Multi-Depth Neural Network and Fusion Deep Neural Networks. The multi-view technique improves diagnosis by combining information from both types of ultrasound images.

2.3.2. Transfer Learning-Based Segmentation Studies

Negi et al. (2020) proposed the WGAN-RDA-UNET technique, which uses Wasserstein GANs. It obtained an

overall accuracy of 0.98 and a PR-AUC of 0.95. The research used the PASCALVOC2012 dataset for training and the Berkeley Segmentation Database (BSDS 300 and BSDS500) for assessment.

Table 1. *Overview of deep learning-based ultrasound screening and diagnosis, DL: deep learning, TL: transfer learning, C=Classification, SE=Segmentation, DE=Detection*

Author		Methods Used	Dataset	Success rate
Jabeen et al. (2022)	DL, C	Modified DarkNet-53 deep model	BUSI 780 images	99.1% accuracy
Liu et al. (2022)	DL, C	Grid-based deep feature generation-based	BUS 600 images	97.18% accuracy
Zhuang et al. (2021)	DL, C	Fuzzy enhancement	BUS 1328 images	95.48% accuracy
Momot et al. (2022)	DL, C	Efficient Net B1	ImageNet dataset	81.26% accuracy
Alrubaie et al. (2023)	DL, C	CNN Inception-v3	A: BUSI 780 images B: BUS 9016	96% accuracy for data A 100% accuracy for data B
Cao et al. (2020)	DL, C	Noise filter network (NF-Net)	BUS	73% accuracy
Kim et al. (2021)	DL, C	Weakly-supervised Deep learning techniques	BUS 1000 images	AUC values of 0.92-0.96
Moon et al. (2020)	DL, C	Dense Net VGG-16 and VGG-Like ResNet	Private 1687 images BUSI: 697 images	91.10% accuracy 94.62% accuracy
Moustafa, et al. (2020)	DL, C	Color Doppler AdaBoost ensemble classifier	159 solid masses	AUC of 0.986
Qi et al. (2019)	DL, C	Automated BC diagnosis model	BUS 8145	87.79% accuracy
Y. Zhang et al. (2021)	TL, C	(CLSTM)	N/A	91% accuracy
Cruz-Ramos et al. (2023)	TL, C	fusion and handcrafted features	BUSI 780 images	ACC of 97.6% PRE of 98%
Pang et al. (2021)	TL, C	Semi-supervised GAN model – TGAN model	BUS 1447	90.41% accuracy
Coronado-Gutiérrez et al. (2019)	TL, C	QUS image analysis techniques	BUS 217 images	86.4% accuracy
Mahoro and Akhloufi (2022)	DL, DE	different screening methods for BC	(RIDER) 1500 BUS 250 images	Dice coefficient: 0.82 Similarity rate: 0.69
Masud et al. (2021)	DL, DE	DenseNet201 - ResNet50	BUS 250 images	100
Li et al. (2022)	DL, DE	BUSnet	BUSI 780 images	100 accuracies
Z. Zhang et al. (2021)	DL, DE	Bayesian YOLOv4	BUSI 21,624 images	Sens: 0.88 FPs/S: 0.19
Negi et al. (2020)	TL, DE	(WGAN) (RDAU-NET)	PASCALVOC2012	Accuracy 0.98
Gómez-Flores and de Albuquerque Pereira (2020)	DL, SE	Multi models	BUS 3000 images	F1-score > 0.90
Ilesanmi et al. (2021)	DL, SE	End-to-end deep learning segmentation stage	Dataset 1: 264 images Dataset 2: 830 images	89.73% were cancerous and 89.62% were benign.
Vakanski et al. (2020)	DL, SE	U-Net - U-Net-SA	BUS 510 images	(DSC) 90.5%
Xu et al. (2019)	DL, SE	CNNs	Private dataset	80% accuracy
Zhang et al. (2019)	DL, SE	BASDL	UDIAT, UTWS	92.00% accuracy for UDIAT 83.9% accuracy for UTWS
Chiang et al. (2018)	DL, SE	3-D CNN - 2-D CNN	Automated whole breast ultrasound images	95% sensitivity
Lei et al. (2018)	DL, SE	ConvEDNet - Adaptive Domain Transfer (ADT)	Automated Whole Breast Ultrasound images	86.8% intersection over union
Gong et al. (2020)	DL, SE	Multi-view deep neural network	BUS 264 images	86.36% accuracy

3. MATERIAL AND METHOD

3.1. Dataset

BUSI is used as part of the validation procedure in this study. The BUSI Dataset includes 780 breast ultrasound pictures categorized as benign, malignant, or normal. Images were taken from 600 female patients between the ages of 25 and 75. The images are 500 by 500 pixels on average and saved in PNG file format. Images were classified into three categories: there are 133 normal photos, 487 benign images, and 210 malignant images. The normal images reveal healthy breasts, whereas the benign images show benign masses and the malignant images indicate cancerous masses. The original images were preprocessed to reduce noise and improve image quality. This involved cropping the images to remove irrelevant information, scaling the images to a uniform size, and applying contrast enhancement. The data was collected in 2018 using LOGIQ E9 ultrasound systems. The images were captured using standard ultrasound protocols for breast imaging. The images were then manually labeled by expert radiologists (Al-Dhabyani et al., 2020). In addition, ground truth the appropriate B-mode images are supplied with binary mask images. Figure 1 shows the ground truth images.

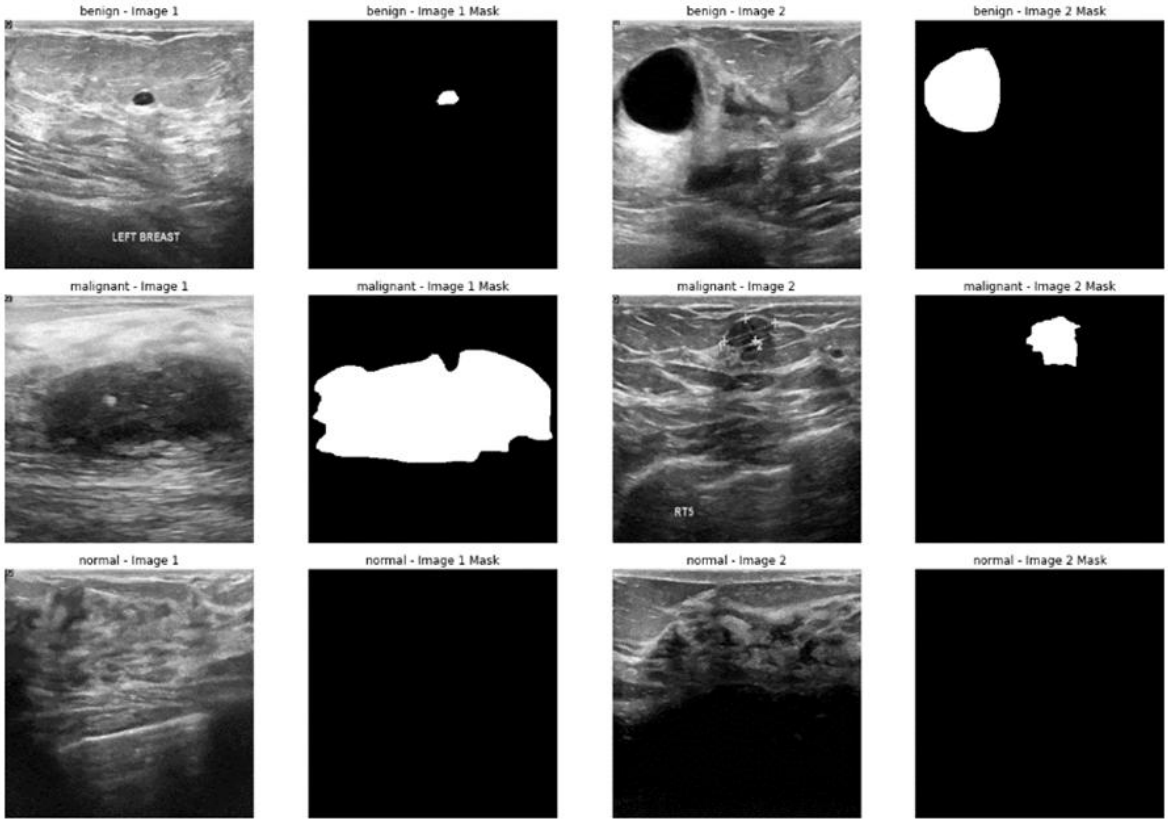


Figure 1. *Examples of ultrasound images with ground truth*

3.2. Methodology

3.2.1. Image Preprocessing and Segmentation

The segmentation section of our workflow is shown in Figure 2. We use a U-Net architecture to precisely identify areas on breast ultrasonography images that are of interest, distinguishing between benign, malignant, and normal tissues. First, we load and preprocess the dataset. This involves resizing the images and their corresponding masks to a uniform 224x224 pixels and normalizing them to ensure consistency. This preparation step is crucial for feeding the data into our model effectively. For medical image segmentation, the U-Net model is especially well-suited. Its architecture comprises two main parts: decreasing and increasing paths. The contracting path (or encoder) captures detailed contextual information by progressively down sampling the images using convolutional layers and max-pooling operations. This step extracts high-level features while reducing the spatial dimensions of the images. At the deepest point, a bottleneck layer further

refines these features. The expanding path (or decoder) then restores the spatial resolution using up sampling layers. It utilizes skip connections to concatenate high-level features from the contracting path with the corresponding layers in the expanding path, effectively recovering spatial information that might have been lost during down sampling. This combination ensures that the output mask accurately highlights the regions of interest. After constructing the model, we compile it using the Adam optimizer and binary cross-entropy loss, with accuracy as the metric to evaluate performance. We divided the dataset into training and validation sets, with 80% for training and 20% for validation. This ensures the model can adequately generalize to new, previously unknown data. Training the U-Net model involves running multiple iterations (epochs) over the dataset, during which the model learns to segment the images accurately. Once training is complete, we save the model for future use. This allows us to efficiently and precisely segment breast lesions in new ultrasound images, which is critical for early and accurate diagnosis of BC. This automated segmentation process significantly enhances the potential for effective treatment and better patient outcomes.

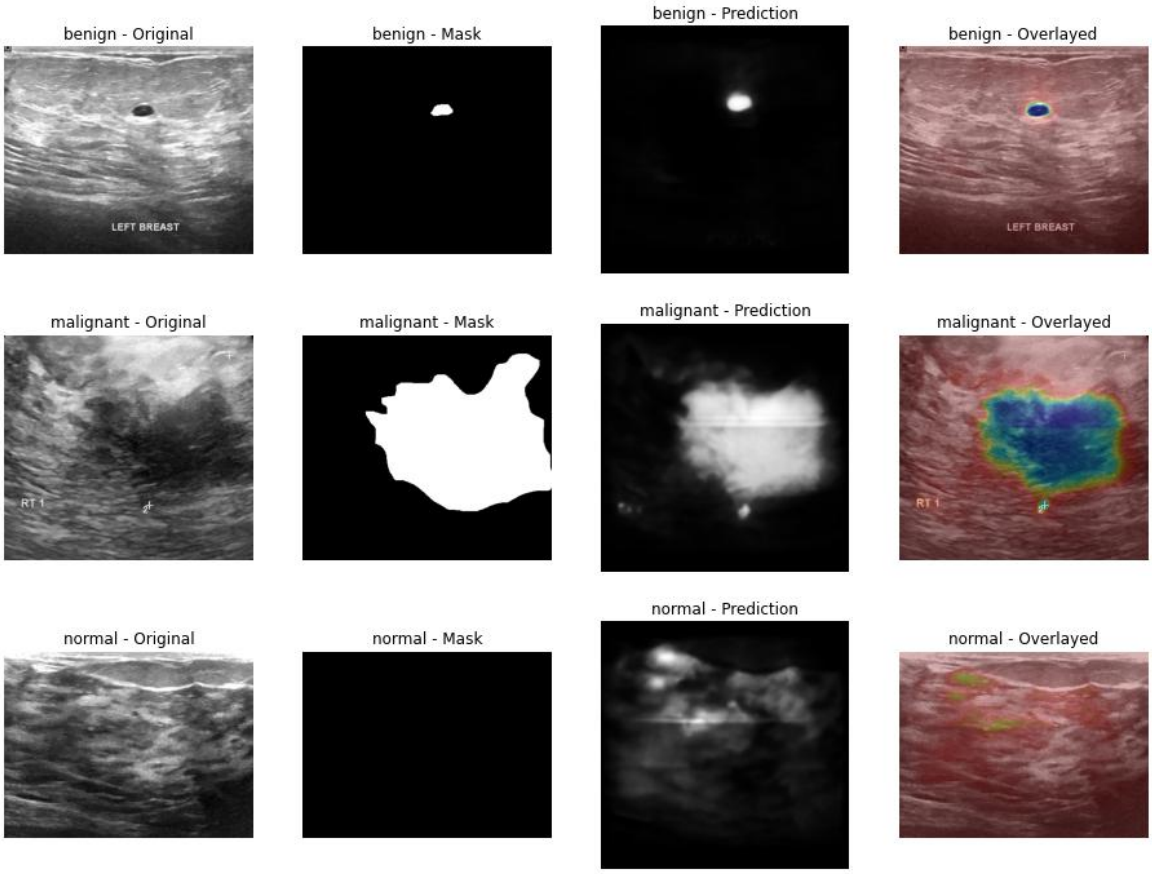


Figure 2. *Visual results for classification and segmentation of breast ultrasound images*

Image enhancement techniques are crucial for medical image analysis as they help improve the visibility of key structures and features. In this context, the BUSI dataset is enhanced using various techniques to make the images more informative and useful for diagnosis. Fuzzy enhancement focuses on adjusting the contrast of an image. By enhancing contrast, subtle differences in tissue density and composition become more apparent, which is crucial for accurate diagnosis in medical imaging. Sharpness enhancement improves the clarity of edges within the image, which is particularly important in ultrasound images where the edges of lesions or tumors need to be well-defined. Adjusting the brightness of an image can help balance the image intensity, which is useful for highlighting regions of interest that may otherwise be obscured due to uneven lighting or intensity variations in the original image. Rescaling images to a uniform size ensures consistency in visualization and comparison, which is essential for processing large datasets and training machine learning models. As shown in Figure 3, the overlay process involves combining the original ultrasound image with the mask that highlights areas of interest, such as tumors or lesions. As shown in Figure 4, this combined image is then enhanced using the above techniques to provide a clearer and more informative visualization.

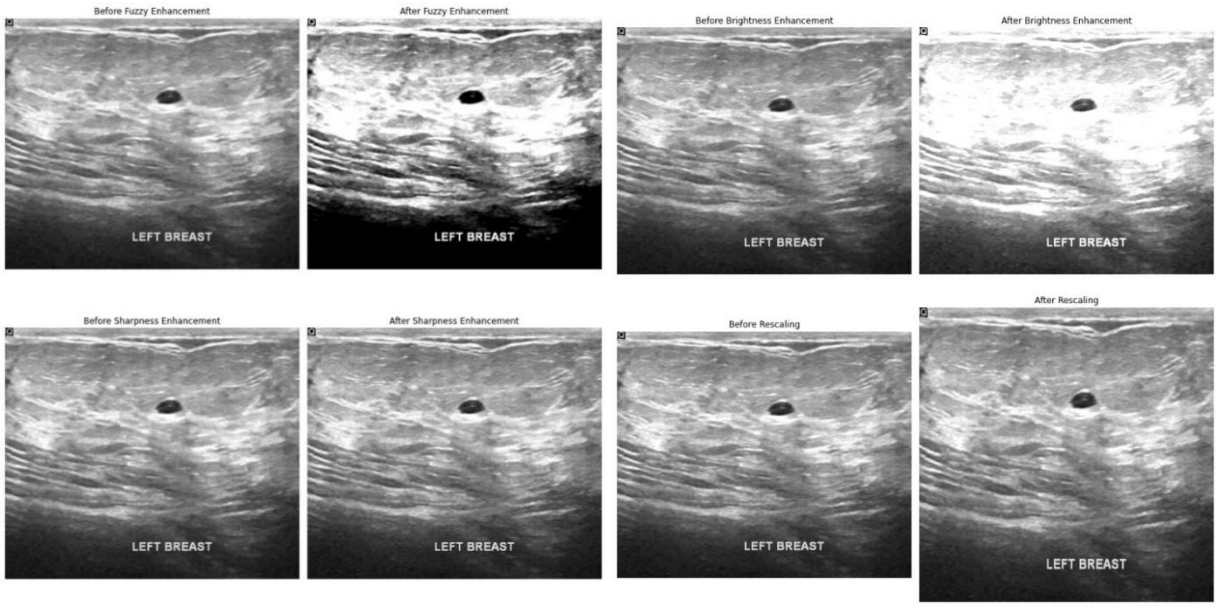
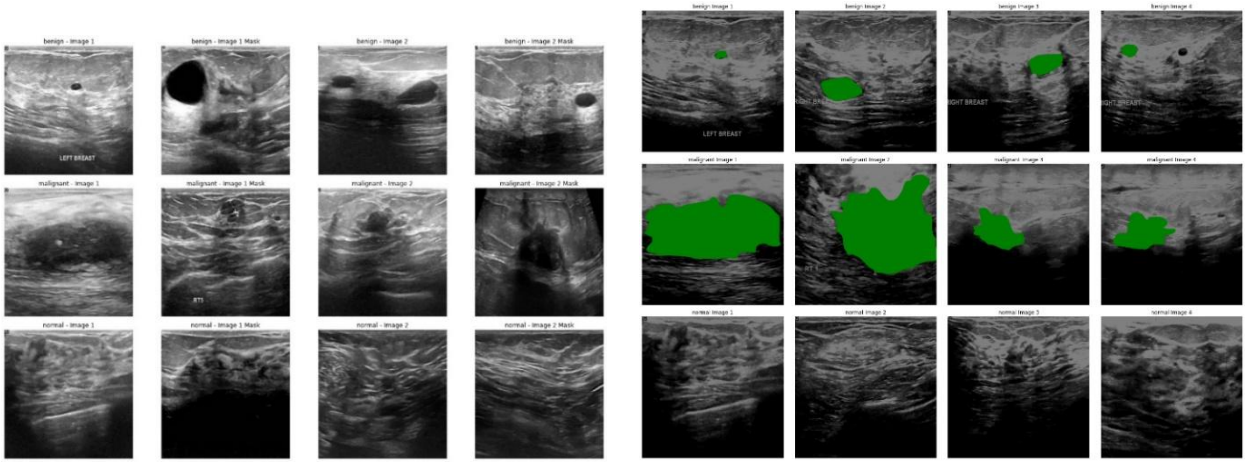


Figure 3. *Ultrasound images of the left breast under different image enhancement techniques*



(a) Before image processing (b) After image processing.

Figure 4. *Enhanced and overlayed image before and after image processing*

3.2.2. Proposed Model

We developed a custom CNN for breast image classification. As shown in Figure 5, our model captures key properties from images using a series of layers of convolution and pooling that preserve the spatial hierarchy. The model starts with a convolutional layer that has 32 filters, each measuring 3x3 pixels. This layer uses a stride of (1, 1) and the same padding to extract key characteristics from the input pictures. It is followed by a max-pooling layer with a kernel size of (2,2) and a stride of (2,2), which reduces spatial dimensions while enhancing translation invariance and model efficiency. Our network has additional convolutional layers with larger filter sizes (64, 128, 256, 256, and 512), followed by max-pooling layers. Except for the last convolutional layer, the output is decreased to 1*1 size to further reduce the sample size of the feature maps while improving abstraction. We use ReLU (Rectified Linear Unit) Functions for activation in these layers to help the network understand complex patterns in the input. The network's depth and structure allow it to captivate both local and global picture functions effectively. Following the convolutional layers, we used GlobalAveragePooling to fully connect the (dense) layer. The model has three dense layers: the first has 512 neurons and employs ReLU activation in conjunction with L2 regularization (at a value of 0.01) to avoid overfitting. We also use a 0.5-rate dropout layer to deactivate neurons during training, randomly increasing generalization. The second layer has 1024 neurons and employs ReLU activation in conjunction with L2

regularization (with a value of 0.01). The last dense layer contains three neurons and employs a softmax activation function to generate probabilistic outputs for the three classes (benign, malignant, and normal). Our Results indicate that the suggested CNN model is highly effective. When we trained it on enriched overlayed datasets, the model had an F1-score of 0.97, with an accuracy of 0.97, precision of 0.98, and recall of 0.96. These metrics were much higher than those obtained while training with all layers without the enhanced overlayed datasets, which yielded an accuracy is 0.83, a precision is 0.80, a recall is 0.83, and an F1-score is 0.81. These findings demonstrate the robustness of our algorithm in detecting subtle patterns in medical images, resulting in extremely accurate categorization. We experimented with various models, architectures, and hyperparameters to achieve an accuracy higher than 97%. However, only this specific model managed to exceed that benchmark on our dataset.

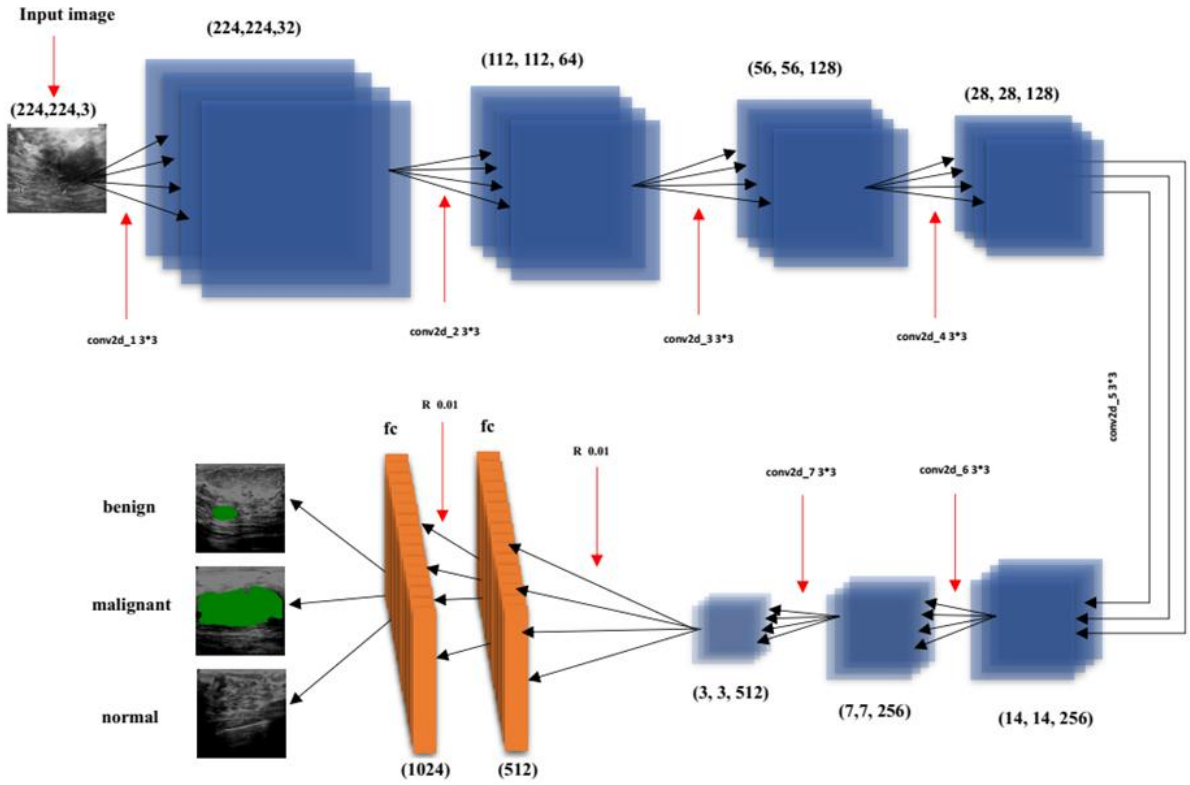


Figure 5. *Proposed model*

3.2.3. State-of-the-art Models

We selected several state-of-the-art models for their effectiveness in enhancing image classification performance: VGG16, VGG19, DenseNet121, EfficientNetB0, MobileNet, InceptionV3, Xception, and ResNet50. VGG16 and VGG19, created by the University of Oxford's Visual Geometry Group, consist of 16 and 19 layers, respectively, with convolutional and max-pooling layers stacked to capture complex features. DenseNet121 employs densely connected blocks, promoting feature reuse and efficient gradient flow. EfficientNetB0 uses compound scaling to optimize model depth, width, and resolution, achieving superior performance. MobileNet offers lightweight architecture suitable for mobile and embedded devices, utilizing depthwise separable convolutions to reduce complexity. InceptionV3 captures features at multiple scales with inception modules, while Xception replaces standard convolutions with depthwise separable convolutions for increased efficiency. ResNet50 solves the issue of vanishing gradients in deep neural networks by utilizing skip connections. During training and assessment, we fine-tuned the last four layers of each pre-trained model using an enhanced dataset. To improve generalization, we employed an Adam optimizer with a 0.0001 learning rate and a sparse categorical cross-entropy loss. We also employed data augmentation techniques including flipping, rotation, shifting, shear, and zooming. The training lasted 100 epochs with a batch size of 32, and it ended early after five epochs to avoid overfitting. Following training, we examined the models on a separate

test dataset, using measures such as accuracy, F1 score, precision, and recall to assess their classification skills. We implemented a custom top-layer architecture tailored for 3-class classification. This architecture builds on a base model, globalAveragePooling2D, which is used to minimize spatial dimensions, followed by dense layers with regularization for dropouts to avoid overfitting. The final output layer, consisting of three neurons with softmax activation, facilitates accurate multi-class classification.

4. RESULTS AND DISCUSSION

4.1. Proposed Model

We found that, when image preprocessing methods were used, the accuracy reached 0.97. Our proposed model performed amazingly with these preprocessing methods regarding other essential metrics: F1-Score: 0.97, precision: 0.98, recall: 0.96. Figure 6 demonstrates accuracy and loss graphs of the proposed model after data preprocessing. When this model was trained without preprocessing, its performance dropped significantly; it still has respectable metrics. More concretely, the accuracy dropped to 83%, meaning the classification would be less accurate than if preprocessing were applied. The confusion matrix elaborated on the classification capabilities of this model, showing it to be proficient at distinguishing classes. Train validation accuracy and loss graphs represented its learning process during training and provided valuable insights into model convergence and generalization. Such findings strongly emphasize the quests of image preprocessing toward improving custom CNN performance but also point to further ways of optimization and research in this field. Regarding accuracy, recall, and F1-score, the proposed model outperformed all other classes. For class 0, the model achieved 97% accuracy, 98% recall, and an F1 score of 97%. For class 1, this results in a 95% accuracy, 94% recall, and a 93% F1 score. Class 2 had 100% accuracy, 100% recall, and an F1 score of 100%. The confusion matrix also presented the model accuracy, showing minimal misclassification. Figure 7 depicts the confusion matrices obtained prior to and after image preprocessing. Overall, these results underscore the effectiveness of the proposed model in accurately classifying instances across multiple classes. Without data enhancement, there was a slight decline in the classification results. Figure 8 represents the accuracy and loss graphs of the proposed model data preprocessing. For class 0, the model recorded a precision of 85%, a recall of 91%, and an F1-score of 100%. For class 1, this gives a precision of 75%, a recall of 86%, and an F1-score of 80%. Class 2 had an accuracy of 68% and a recall of 56%, with an F1 score of 71%. Despite these variations, the confusion matrix showed the model's accuracy, with minimal misclassifications. These findings highlight the proposed model's effectiveness in accurately classifying instances across multiple classes while demonstrating the significant impact of data preprocessing on model performance.

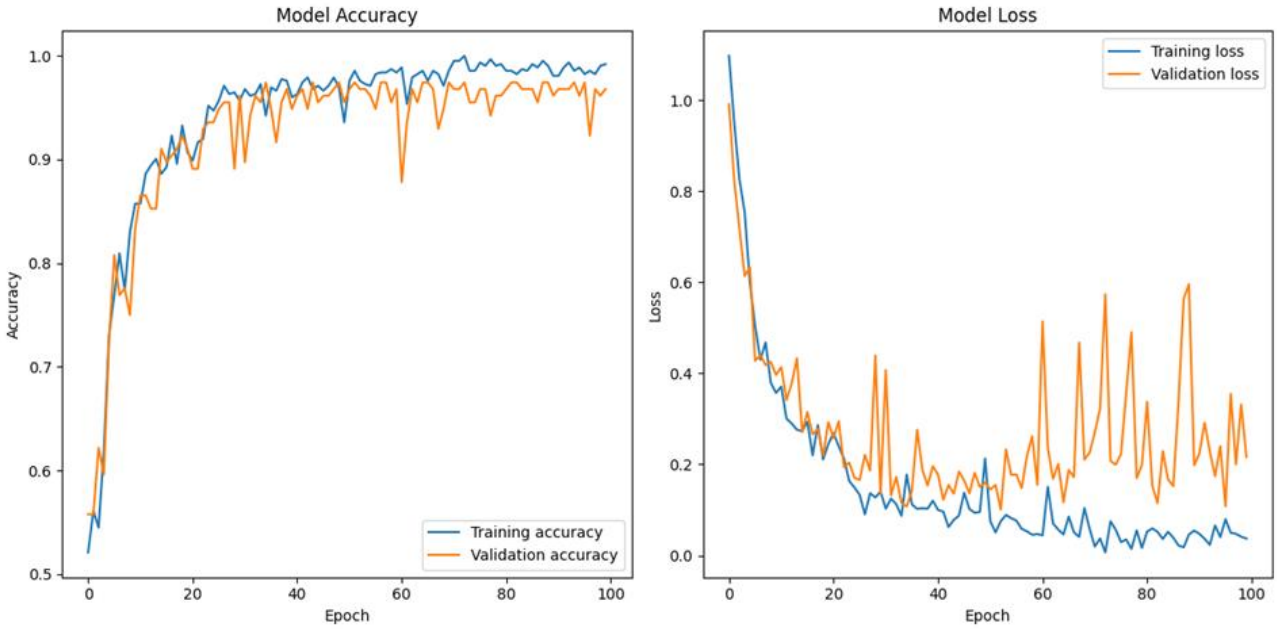


Figure 6. *The accuracy and the loss of the proposed model after data preprocessing*

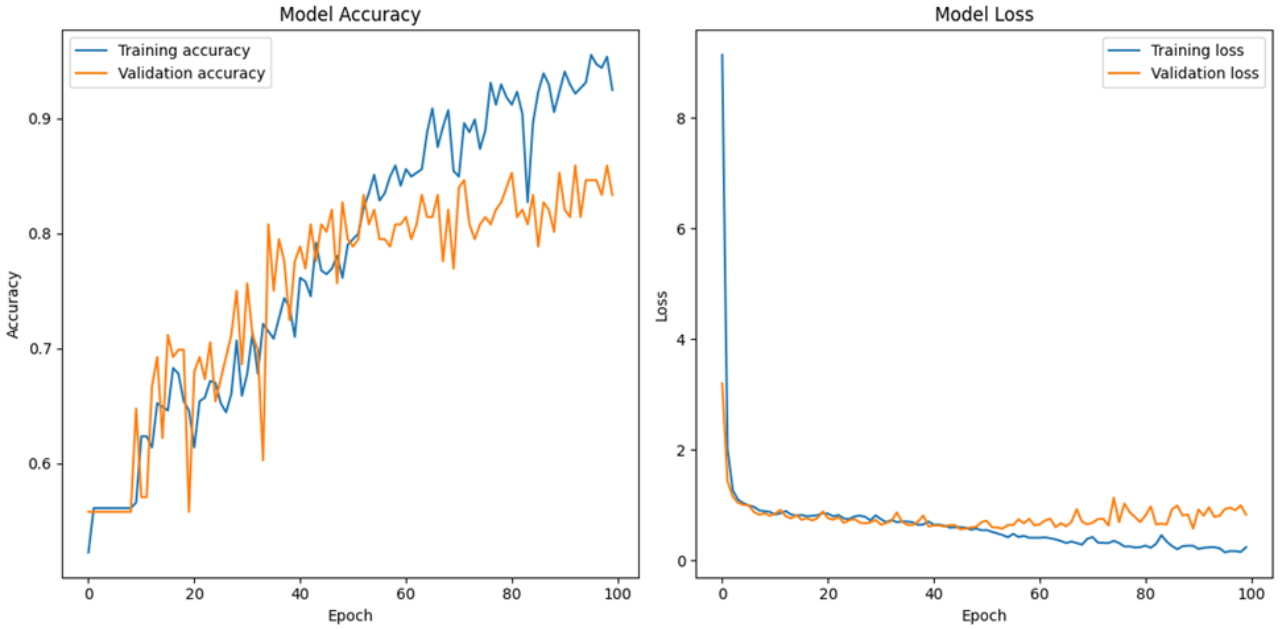
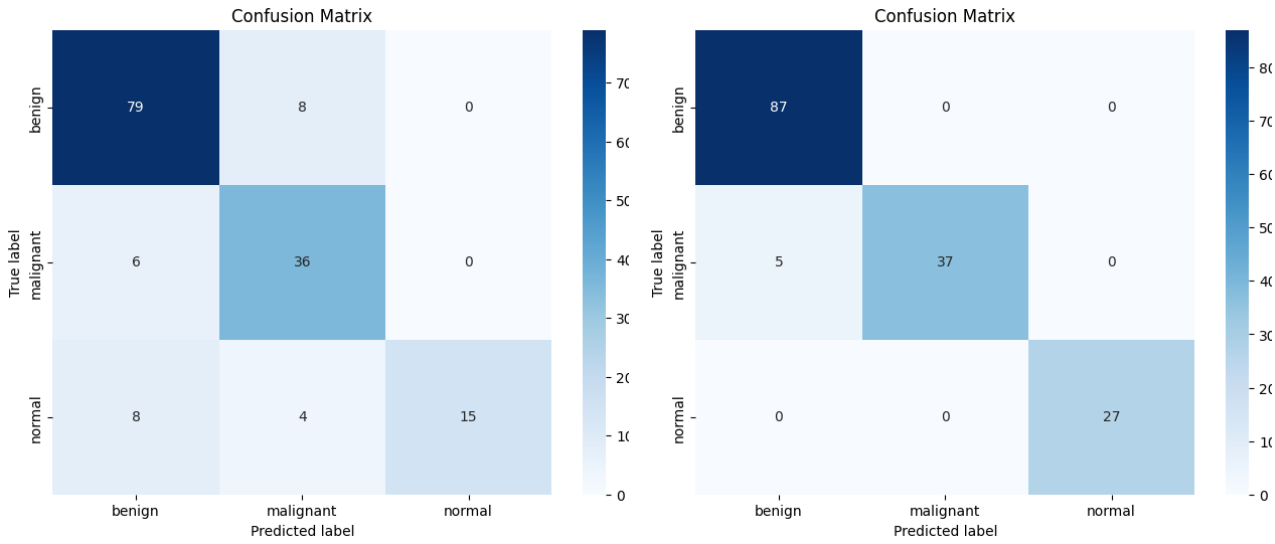


Figure 7. *The accuracy and the loss of the proposed model before data preprocessing*



(a) Before image processing (b) After image processing

Figure 8. *Confusion matrix of the proposed model before and after image preprocessing*

4.2. Transfer Learning

In this study, we investigated the performance of fine-tuned models on a dataset augmented with enhanced overlayed images, considering various training scenarios. The performance metrics for each scenario are summarized in Tables 2, 3, 4, and 5. Table 2 presents the performance of models fine-tuned with the last four layers using the enhanced overlayed dataset. VGG19 and EfficientNetB0 demonstrated the highest accuracies of 0.99, with VGG19 achieving a precision of 0.98, a recall of 0.98, and an F1-score of 0.98. VGG16 and ResNet50 also performed exceptionally well, with accuracies of 0.98 and 0.96, respectively. This demonstrates the effectiveness of the enhanced dataset in improving model performance, achieving balanced and high scores across all metrics.

Table 2. *Performance metrics of the models trained with the last four layers with enhancing and overlaying*

Model	Accuracy	Precision	Recall	F1-Score
VGG16	0.98	0.99	0.98	0.98
VGG19	0.99	0.98	0.98	0.98
DenseNet121	0.88	0.91	0.87	0.89
EfficientNetB0	0.97	0.99	0.97	0.98
MobileNet	0.82	0.82	0.81	0.80
Inceptionv3	0.63	0.59	0.55	0.56
Xception	0.69	0.67	0.61	0.63
ResNet50	0.96	0.96	0.96	0.96

Table 3 shows the performance of models fine-tuned with the last four layers without enhancements. There was a noticeable drop in performance across all models. VGG16's accuracy decreased to 0.82, and Inceptionv3's accuracy fell to 0.60. EfficientNetB0 and VGG19, while still relatively higher performing with accuracies of 0.86 and 0.84, respectively, also exhibited declines in their precision, recall, and F1 scores. This indicates the negative impact of the absence of enhancements on model generalization and performance.

Table 3. *Performance metrics of the models trained with the last four layers without enhancing and overlaying*

Model	Accuracy	Precision	Recall	F1-Score
VGG16	0.82	0.79	0.82	0.80
VGG19	0.84	0.81	0.85	0.82
DenseNet121	0.76	0.75	0.71	0.72
EfficientNetB0	0.86	0.85	0.84	0.84
MobileNet	0.74	0.69	0.72	0.70
Inceptionv3	0.60	0.57	0.50	0.51
Xception	0.66	0.61	0.58	0.59
ResNet50	0.81	0.83	0.75	0.78

Table 4 represents the results of models trained with all layers using the enhanced overlayed dataset. These models exhibited outstanding performance, with several achieving near-perfect metrics. EfficientNetB0 reached an accuracy of 0.99, with precision, recall, and an F1-score of around 0.99. VGG19 and DenseNet121 both attained an accuracy of 0.98. The proposed model also showed strong performance with an accuracy of 0.97. These results underscore the benefits of using an enhanced dataset for training models across all layers.

Table 4. *Performance metrics of the models trained with all layers with enhancing and overlaying*

Model	Accuracy	Precision	Recall	F1-Score
VGG16	0.96	0.98	0.95	0.96
VGG19	0.98	0.99	0.98	0.98
DenseNet121	0.98	0.99	0.98	0.98
EfficientNetB0	0.99	0.99	0.98	0.99
MobileNet	0.98	0.98	0.98	0.98
Inceptionv3	0.98	0.99	0.98	0.98
Xception	0.98	0.99	0.98	0.98
ResNet50	0.91	0.91	0.93	0.92
Proposed Model	0.97	0.97	0.97	0.97

Table 5 details the performance of models trained with all layers without enhancements. Although there was a general decrease in performance compared to their enhanced counterparts, models like DenseNet121 and EfficientNetB0 maintained relatively high accuracies of 0.87 and 0.86, respectively. However, models such as VGG16 and MobileNet exhibited more substantial drops, emphasizing the importance of dataset enhancements. The proposed model, without improvements, showed a moderate decrease in performance with an accuracy of 0.83. Our study demonstrates the critical role of dataset quality and preprocessing techniques in enhancing deep learning models' robustness and generalization capabilities. The enhanced overlayed dataset consistently led to significant improvements in model accuracy, precision, recall, and F1 score across various architectures. These findings confirm the value of data enhancement in machine learning tasks, highlighting its importance for achieving superior model performance. Figure 9 and Figure 10 represent the confusion matrices of the transfer learning models after and before data preprocessing.

Table 5. *Performance metrics of the models trained with all layers without enhancing and overlaying*

Model	Accuracy	Precision	Recall	F1-Score
VGG16	0.76	0.75	0.73	0.71
VGG19	0.77	0.74	0.77	0.75
DenseNet121	0.87	0.89	0.83	0.85
EfficientNetB0	0.86	0.85	0.84	0.85
MobileNet	0.86	0.85	0.84	0.83
Inceptionv3	0.82	0.80	0.80	0.80
Xception	0.85	0.84	0.81	0.82
ResNet50	0.85	0.84	0.79	0.81
Proposed Model	0.83	0.87	0.77	0.80

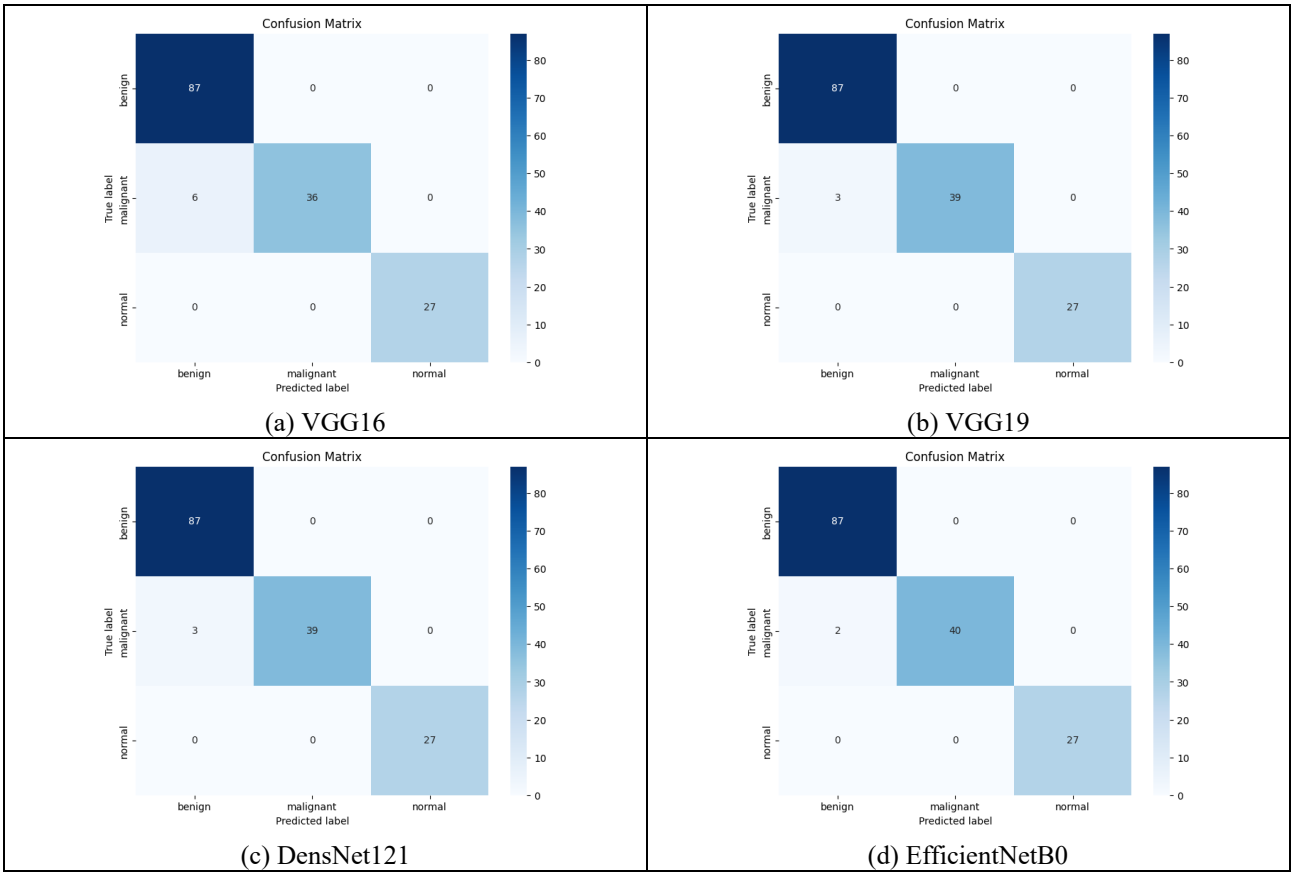


Figure 9. *Confusion matrix of transfer learning models after data preprocessing*

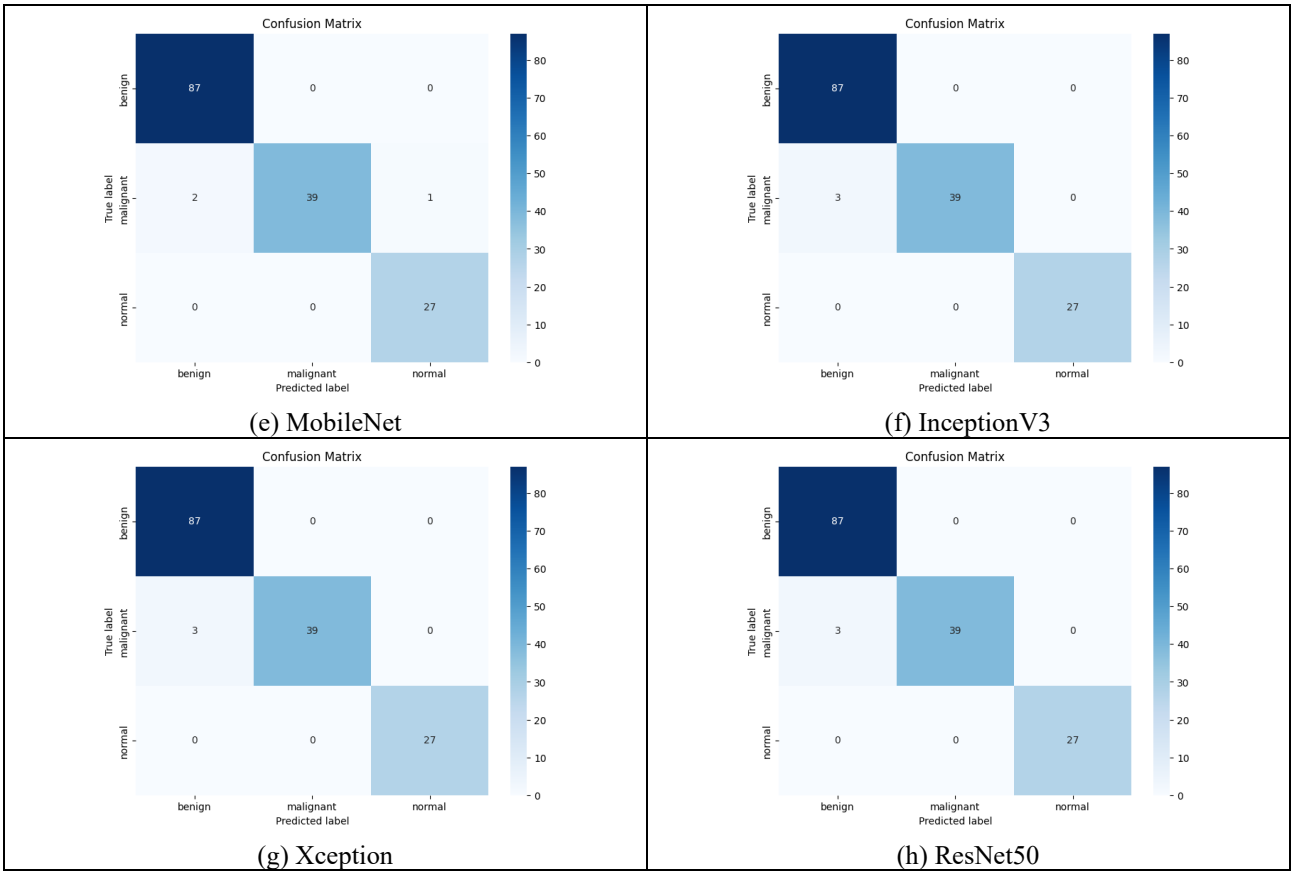


Figure 9. *continued*

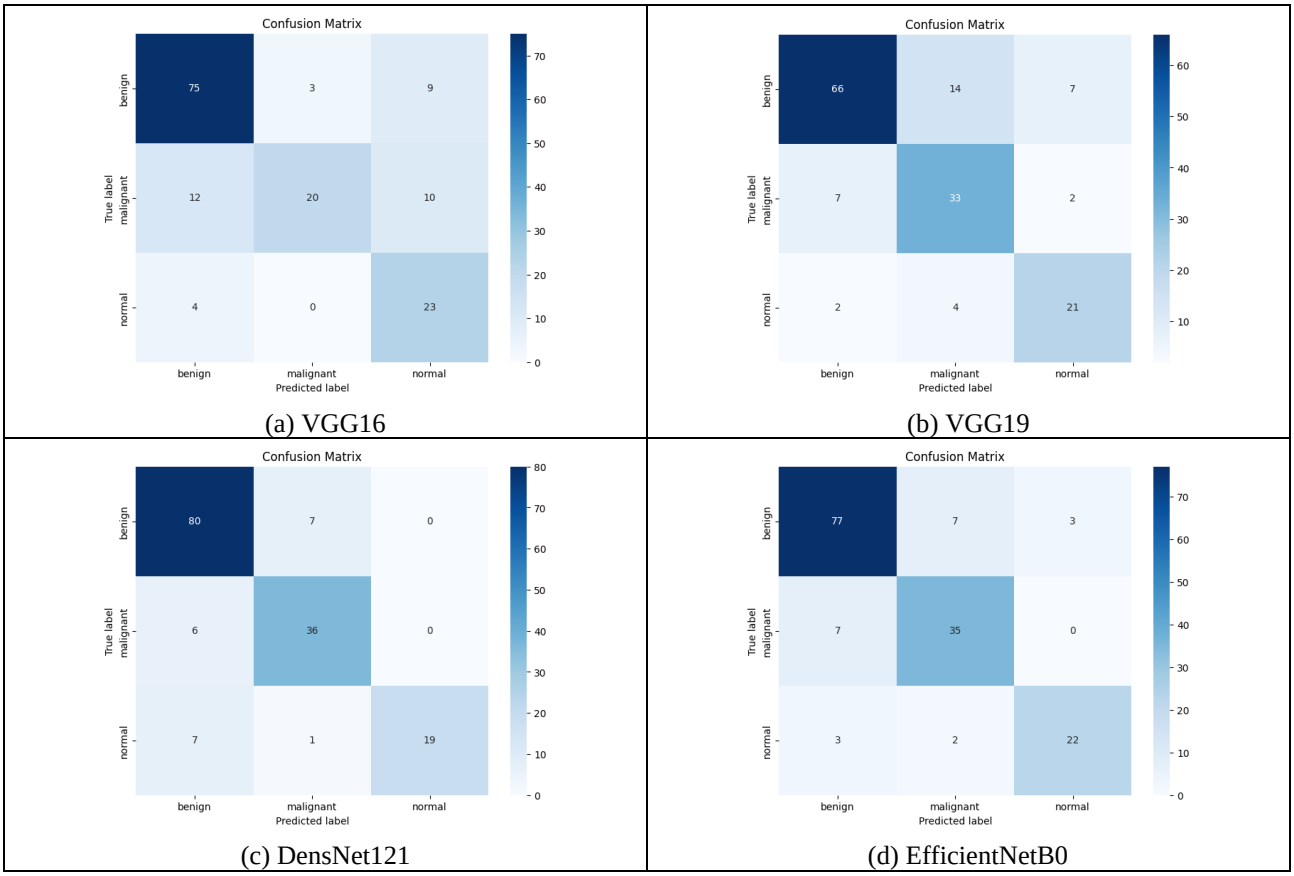
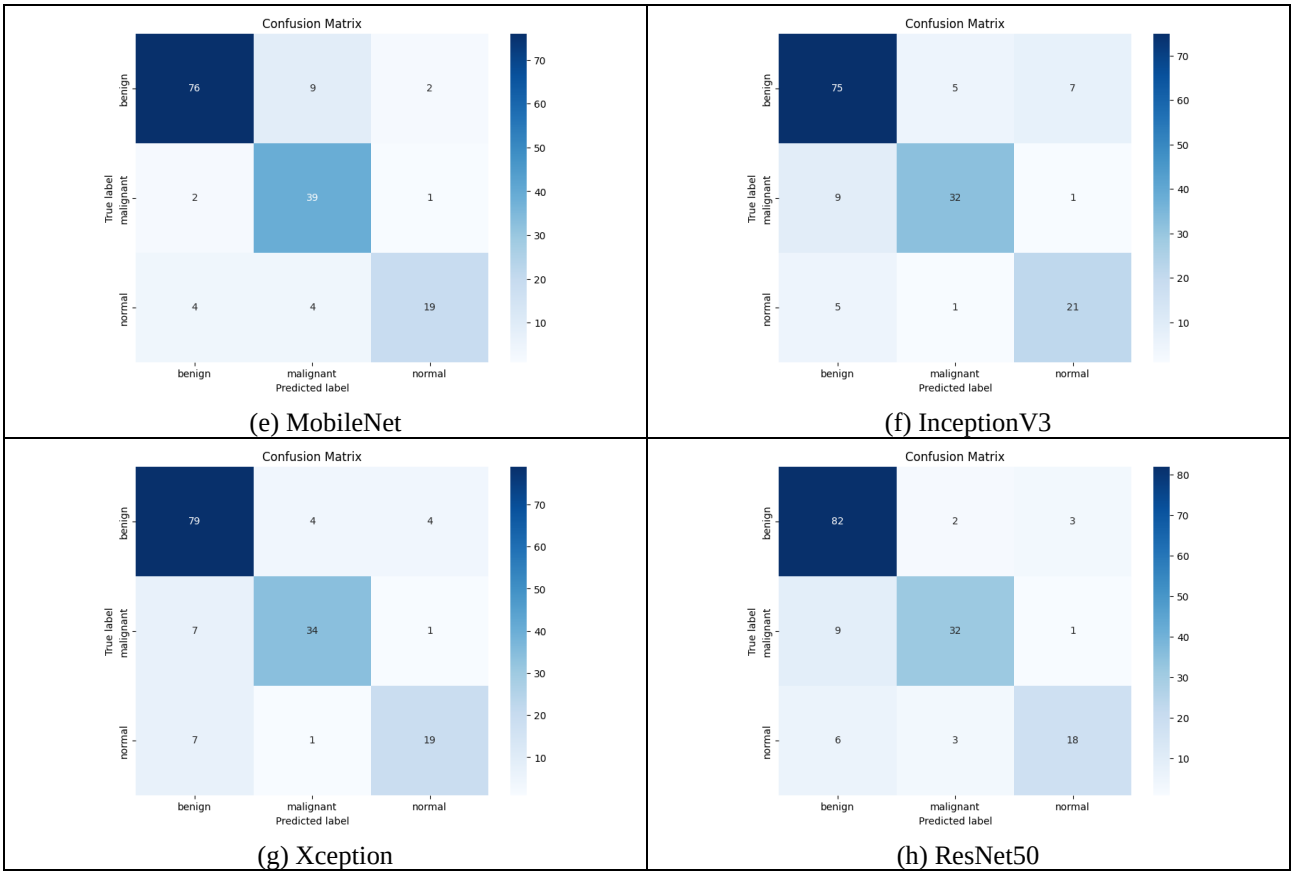


Figure 10. *Confusion matrix of transfer learning models before data preprocessing*



(e) MobileNet (f) InceptionV3

(g) Xception (h) ResNet50

Figure 10. *continued*

4.3. Comparison of the Models

When comparing our study's findings with other transfer learning models, EfficientNetB0 performed exceptionally well, achieving an accuracy score of 0.99. This indicates its high confidence and precision in categorizing images across various datasets. VGG19, MobileNet, DenseNet121, InceptionV3, and Xception, all performed well with an accuracy of 0.98. VGG16 achieved an accuracy of 0.96. ResNet50, although slightly lower, still maintained strong performance with an accuracy of 0.91. The proposed model also demonstrated excellent performance, achieving an accuracy of 0.97. The proposed model achieved high precision, recall, and F1 scores with significantly fewer parameters than other architectures except EfficientNetB0 and Mobilenet. VGG16 with 14879299, VGG19 with 20188995, Densnet121 with 7333187, EfficientNetB0 with 4410790, Mobilenet with 3524547, InceptionV3 with 22360611, Xception with 21419307, ResNet50 with 24145539. With only 3097283 parameters, our model efficiently extracts relevant features and makes precise classifications without needing an excessive number of parameters. Figure 11 shows the comparison of transfer learning and our proposed model. This emphasizes not only its technical excellence but also its efficiency in using computational resources. Therefore, our proposed model emerges as a standout performer, offering a strong case for its adoption in image classification tasks where both efficiency and accuracy are crucial.

4.4. Comparison with Similar Studies

Several important insights emerge when comparing our study's outcomes with similar studies utilizing deep learning or transfer learning methodologies and comparable datasets for BC analysis. Liu et al. (2022) followed suit with a grid-based deep feature generation approach on the BUS dataset, yielding a commendable accuracy of 97.18%, showcasing the efficacy of this method in achieving high precision. Zhuang et al. (2021) employed fuzzy enhancement techniques on the BUS dataset, resulting in a notable accuracy of 95.48%, demonstrating the potential of such preprocessing methods in improving classification performance. Alrubaie et al. (2023) utilized the Inception-v3 CNN architecture on both BUSI and BUS datasets, achieving remarkable accuracies of 96% and 100%, respectively, underscoring the robustness of their approach across different datasets. Cruz-

Ramos et al. (2023) integrated fusion and handcrafted features on the BUSI dataset, attaining an accuracy of 97.6% and a precision of 98%, showing the effectiveness of feature engineering in enhancing classification outcomes. Pang et al. (2021) leveraged a semi-supervised GAN model on the BUS dataset, achieving a notable accuracy of 90.41%, highlighting the potential of generative models in augmenting classification tasks. Z. Zhang et al. (2021) employed Bayesian YOLOv4 on a large-scale BUSI dataset, achieving a sensitivity of 0.88 and a low rate of false positives, demonstrating the efficacy of their approach in achieving accurate tumor detection. Compared to these studies, our approach, which utilized a custom CNN architecture tailored specifically for breast ultrasound analysis, achieved competitive accuracy. We achieved 97% accuracy, demonstrating the effectiveness of our methodology for accurately classifying breast ultrasound. Through this comparative analysis, we gain valuable insights into the diverse methods and their respective successes in BC analysis tasks. Table 6 shows the comparison of our model with related studies.

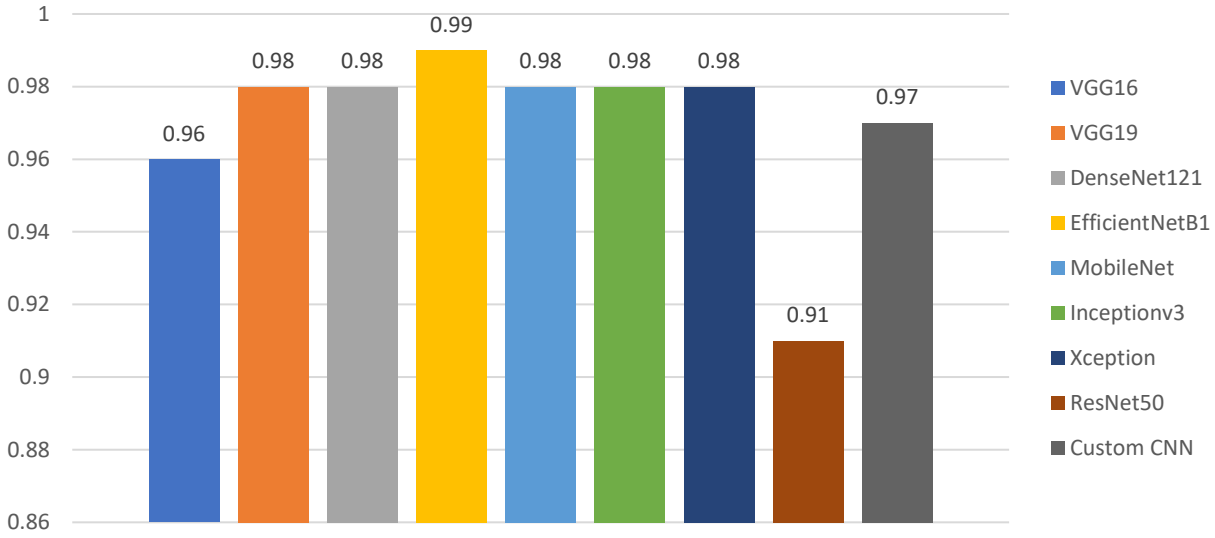


Figure 11. *Comparison chart of transfer learning and Custom CNN (proposed model)*

Table 6. *Comparison table of Proposed Model with related studies*

Study	Methods Used	Dataset	Success Rate
Zhuang et al. (2021)	Fuzzy enhancement	BUS (1328 images)	95.48%
Cruz-Ramos et al. (2023)	Fusion and handcrafted features	BUSI (780 images)	97.6%
Pang et al. (2021)	Semi-supervised GAN model	BUS (1447 images)	90.41%
Z. Zhang et al. (2021)	Bayesian YOLOv4	BUSI (21,624 images)	Sens: 0.88 FPs/S: 0.19
Alrubaie et al. (2023)	Inception-v3 CNN	A: BUSI (780 images) B: BUS (9016 images)	96%
Proposed Model	CNN	BUSI (780 images)	97%

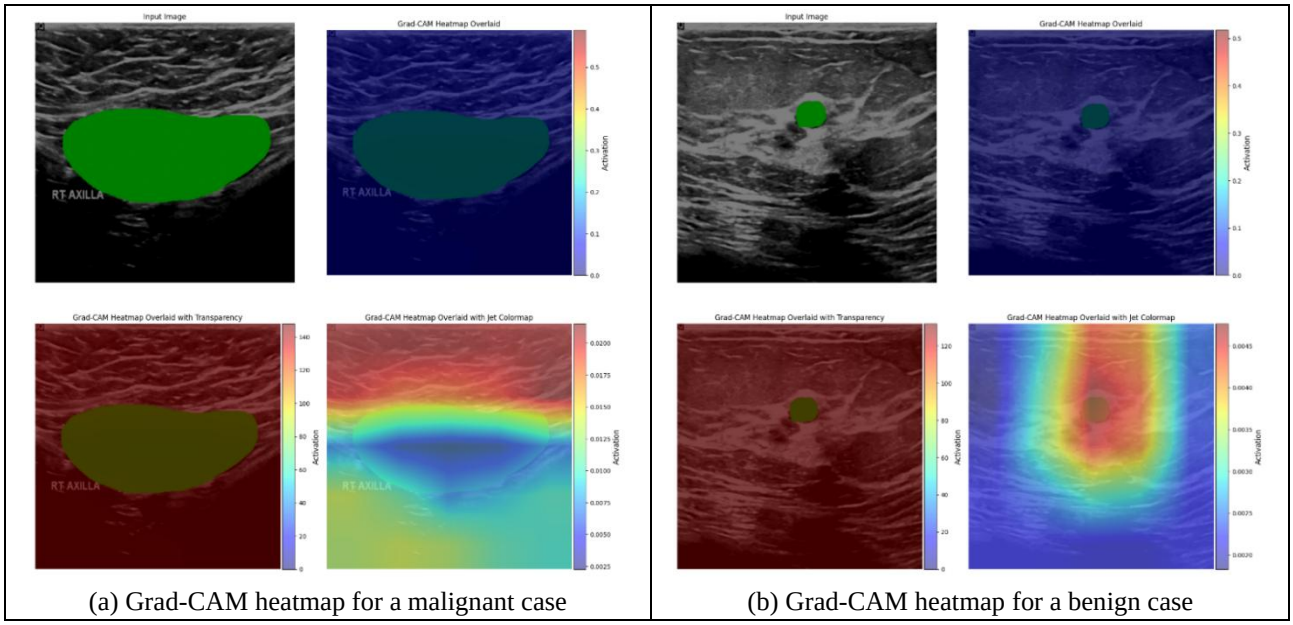
4.5. The Grad-CAM Visualizing

The Grad-CAM (Gradient-weighted Class Activation Mapping) technique was utilized to understand the internal processes of a CNN model. At the same time, it made predictions on images from the test dataset. By analyzing the activations of the final convolutional layer of the CNN, Grad-CAM generates heatmaps that highlight the regions of the input image that are most influential in determining the model's output. In this analysis, specific layers of the CNN architecture were chosen to visualize the activation patterns corresponding to different classes. Figure 12 presents the Grad-CAM heatmap for a malignant case, the heatmap for a benign case, and the heatmap for a normal case. Each figure exhibits four distinct visualizations, providing a comprehensive understanding of the model's decision-making process. First, the original input image offers

context. Subsequently, the Grad-CAM heatmap is overlaid onto the input image, offering a visual representation of the areas within the image that strongly influenced the model's prediction. To further enhance interpretability, two additional visualizations were created: one with the Grad-CAM heatmap overlaid with transparency, providing a clearer view of the heatmap's distribution, and the other utilizing a transparent jet colormap, facilitating a more intuitive interpretation of the heatmap. These visualizations not only aid in understanding which image regions contribute most to the model's decision but also offer valuable insights into CNN's feature extraction and classification processes. By clarifying the model's attention mechanism, Grad-CAM enables researchers and practitioners to validate and interpret the model's predictions, enhancing deep learning models' transparency, trustworthiness, and interpretability in image analysis tasks.

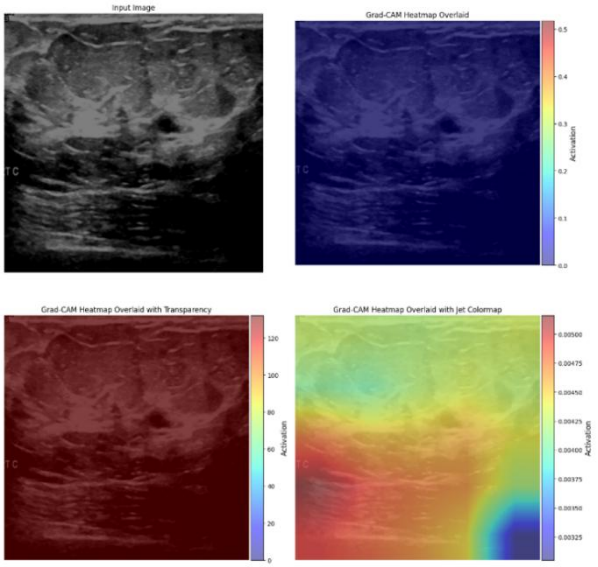
5. CONCLUSION

In conclusion, this research demonstrated the enormous potential of deep learning models, especially CNNs, for the efficient classification of ultrasound images in the diagnosis of BC. The subject models were easily interpretable; besides, they showed better diagnostic performance when advanced image processing techniques-segmentation and enhancement-were applied to the images. Our proposed CNN model achieved an accuracy of 97%, compared with other state-of-the-art models such as EfficientNetB0, MobileNet, and InceptionV3. This emphasizes the need for models specially tailored with the ability to handle challenges that come with the uniqueness of medical imaging tasks. More importantly, the interpretability and transparency of our models were taken further with Grad-CAM use, making the decision-making process more comprehensible. This is critical in the clinical field. AI diagnostic tools should be not only precise but also interpretable for healthcare professionals and patients to build confidence in them. These results also suggest the importance of strong techniques for image preprocessing. Techniques such as image enhancement and overlay greatly improved model performance. This work should be extended in further studies to include more data modalities, such as 3D ultrasound or MRI images, and more complex CNN architectures that may result in even better classification outcomes. Moreover, such models' ability to generalize on larger datasets with more diverse data would be reassuring against their clinical use. The current study provided a sufficient backbone regarding the application of CNNs in the diagnosis of BC by using ultrasound images and gave important indications on how to design more efficient and trustworthy diagnostic tools that could potentially improve patient outcomes.



(a) Grad-CAM heatmap for a malignant case (b) Grad-CAM heatmap for a benign case

Figure 12. *The Grad-CAM heatmap for a malignant case, the heatmap for a benign case, and the heatmap for a normal case*



(c) Grad-CAM heatmap for a normal case

Figure 12. *continued*

AUTHOR CONTRIBUTIONS

Conceptualization, J.C. and Y.C-K.; methodology, J.C. and Y.C-K.; experiments, J.C.; validation, J. C. and Y.C-K.; manuscript-original draft, J.C. and Y.C-K.; manuscript-review and editing, J.C. and Y.C-K; visualization, J.C.; supervision, Y.C-K.. All authors have read and legally accepted the final version of the article published in the journal.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Al-Dhabyani, W., Gomaa, M., Khaled, H., & Fahmy, A. (2020). Dataset of breast ultrasound images. *Data in Brief*, *28*, 104863. https://doi.org/10.1016/j.dib.2019.104863

Alrubaie, H., Aljobouri, H. K., AL-Jobawi, Z. J., & Çankaya, I. (2023). Convolutional neural network deep learning model for improved ultrasound breast tumor classification. *Al-Nahrain Journal for Engineering Sciences*, *26*(2), 57-62. https://doi.org/10.29194/NJES.26020057

Badawy, S. M., Mohamed, A. E.-N. A., Hefnawy, A. A., Zidan, H. E., GadAllah, M. T., & El-Banby, G. M. (2021). Automatic semantic segmentation of breast tumors in ultrasound images based on combining fuzzy logic and deep learning: A feasibility study. *PLOS ONE*, *16*(5), e0251899. https://doi.org/10.1371/journal.pone.0251899

Cao, Z., Yang, G., Chen, Q., Chen, X., & Lv, F. (2020). Breast tumor classification through learning from noisy labeled ultrasound images. *Medical Physics*, *47*(3), 1048-1057. https://doi.org/10.1002/mp.13966

Chiang, T.-C., Huang, Y.-S., Chen, R.-T., Huang, C.-S., & Chang, R.-F. (2018). Tumor detection in automated breast ultrasound using 3-D CNN and prioritized candidate aggregation. *IEEE Transactions on Medical Imaging*, *38*(1), 240-249. https://doi.org/10.1109/TMI.2018.2860257

Coronado-Gutiérrez, D., Santamaría, G., Ganau, S., Bargalló, X., Orlando, S., Oliva-Brañas, M. E., Perez-Moreno, A., & Burgos-Artizzu, X. P. (2019). Quantitative ultrasound image analysis of axillary lymph nodes to diagnose metastatic involvement in breast cancer. *Ultrasound in Medicine & Biology*, *45*(11), 2932-2941. https://doi.org/10.1016/j.ultrasmedbio.2019.07.413

Cruz-Ramos, C., García-Ávila, O., Almaraz-Damián, J.-A., Ponomaryov, V., Reyes-Reyes, R., & Sadovnychiy, S. (2023). Benign and malignant breast tumor classification in ultrasound and mammography

images via fusion of deep learning and handcraft features. *Entropy*, *25*(7), 991. https://doi.org/10.3390/e25070991

Çetin-Kaya, Y., & Kaya, M. (2024). A novel ensemble framework for multi-classification of brain tumors using magnetic resonance imaging. *Diagnostics*, *14*(4), 383. https://doi.org/10.3390/diagnostics14040383

Fujioka, T., Mori, M., Kubota, K., Oyama, J., Yamaga, E., Yashima, Y., Katsuta, L., Nomura, K., Nara, M., Oda, G., Nakagawa, T., Kitazume, Y., & Tateishi, U. (2020). The utility of deep learning in breast ultrasonic imaging: A review. *Diagnostics*, *10*(12), 1055. https://doi.org/10.3390/diagnostics10121055

Gómez-Flores, W., & de Albuquerque Pereira, W. C. (2020). A comparative study of pre-trained convolutional neural networks for semantic segmentation of breast tumors in ultrasound. *Computers in Biology and Medicine*, *126*, 104036. https://doi.org/10.1016/j.compbiomed.2020.104036

Gong, B., Shen, L., Chang, C., Zhou, S., Zhou, W., Li, S., & Shi, J. (2020, April 3-7). *BI-modal ultrasound breast cancer diagnosis via multi-view deep neural network SVM*. In: Proceedings of the 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI) (pp. 1106-1110), Iowa City, IA, USA. IEEE. https://doi.org/10.1109/ISBI45749.2020.9098438

Han, S., Kang, H.-K., Jeong, J.-Y., Park, M.-H., Kim, W., Bang, W.-C., & Seong, Y.-K. (2017). A deep learning framework for supporting the classification of breast lesions in ultrasound images. *Physics in Medicine & Biology*, *62*(19), 7714. https://doi.org/10.1088/1361-6560/aa82ec

Ilesanmi, A. E., Chaumrattanakul, U., & Makhanov, S. S. (2021). A method for segmentation of tumors in breast ultrasound images using the variant enhanced deep learning. *Biocybernetics and Biomedical Engineering*, *41*(2), 802-818. https://doi.org/10.1016/j.bbe.2021.05.007

Jabeen, K., Khan, M. A., Alhaisoni, M., Tariq, U., Zhang, Y.-D., Hamza, A., Mickus, A., & Damaševičius, R. (2022). Breast cancer classification from ultrasound images using probability-based optimal deep learning feature fusion. *Sensors*, *22*(3), 807. https://doi.org/10.3390/s22030807

Kabir, S. M., Bhuiyan, M. I. H., Tanveer, M. S., & Shihavuddin, ASM. (2021). RiIG modeled WCP image-based CNN architecture and feature-based approach in breast tumor classification from B-mode ultrasound. *Applied Sciences*, *11*(24), 12138. https://doi.org/10.3390/app112412138

Kaya, M., & Çetin-Kaya, Y. (2024). A novel ensemble learning framework based on a genetic algorithm for the classification of pneumonia. *Engineering Applications of Artificial Intelligence*, *133*, 108494. https://doi.org/10.1016/j.engappai.2024.108494

Kim, J., Kim, H. J., Kim, C., Lee, J. H., Kim, K. W., Park, Y. M., Kim, H. W., Ki, S. Y., Kim, Y. M., & Kim, W. H. (2021). Weakly supervised deep learning for ultrasound diagnosis of breast cancer. *Scientific Reports*, *11*(1), 24382. https://doi.org/10.1038/s41598-021-03806-7

Lei, B., Huang, S., Li, R., Bian, C., Li, H., Chou, Y.-H., & Cheng, J.-Z. (2018). Segmentation of breast anatomy for automated whole breast ultrasound images with boundary regularized convolutional encoder-decoder network. *Neurocomputing*, *321*, 178-186. https://doi.org/10.1016/j.neucom.2018.09.043

Li, Y., Gu, H., Wang, H., Qin, P., & Wang, J. (2022). BUSnet: A deep learning model of breast tumor lesion detection for ultrasound images. *Frontiers in Oncology*, *12*, 848271. https://doi.org/10.3389/fonc.2022.848271

Liu, H., Cui, G., Luo, Y., Guo, Y., Zhao, L., Wang, Y., Subasi, A., Dogan, S., & Tuncer, T. (2022). Artificial intelligence-based breast cancer diagnosis using ultrasound images and grid-based deep feature generator. *International Journal of General Medicine*, *15*, 2271-2282. https://doi.org/10.2147/IJGM.S347491

Luo, L., Wang, X., Lin, Y., Ma, X., Tan, A., Chan, R., Vardhanabhuti, V., Chu, W. C., Cheng, K.-T., & Chen, H. (2024). Deep learning in breast cancer imaging: A decade of progress and future directions. *IEEE Reviews in Biomedical Engineering*. https://doi.org/10.1109/RBME.2024.3357877

Mahoro, E., & Akhloufi, M. A. (2022). Applying deep learning for breast cancer detection in radiology. *Current Oncology*, *29*(11), 8767-8793. https://doi.org/10.3390/curroncol29110690

Marini, T. J., Castaneda, B., Iyer, R., Baran, T. M., Nemer, O., Dozier, A. M., Parker, K. J., Zhao, Y., Serratelli, W., Matos, G., Ali, S., Ghobryal, B., Visca, A., & O'Connell, A. (2023). Breast ultrasound volume sweep

imaging: A new horizon in expanding imaging access for breast cancer detection. *Journal of Ultrasound in Medicine*, *42*(4), 817-832. https://doi.org/10.1002/jum.16047

Masud, M., Hossain, M. S., Alhumyani, H., Alshamrani, S. S., Cheikhrouhou, O., Ibrahim, S., Muhammad, G., Rashed, A. E. E., & Gupta, B. B. (2021). Pre-trained convolutional neural networks for breast cancer detection using ultrasound images. *ACM Transactions on Internet Technology*, *21*(4), 1-17. https://doi.org/10.1145/3418355

Momot, A., Galagan, R., & Zaboluieva, M. (2022). Automation of ultrasound breast cancer image classification using deep neural networks. *Sciences of Europe*, (96), 38-41.

Moon, W. K., Lee, Y.-W., Ke, H.-H., Lee, S. H., Huang, C.-S., & Chang, R.-F. (2020). Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks. *Computer Methods and Programs in Biomedicine*, *190*, 105361. https://doi.org/10.1016/j.cmpb.2020.105361

Moustafa, A. F., Cary, T. W., Sultan, L. R., Schultz, S. M., Conant, E. F., Venkatesh, S. S., & Sehgal, C. M. (2020). Color Doppler ultrasound improves machine learning diagnosis of breast cancer. *Diagnostics*, *10*(9), 631. https://doi.org/10.3390/diagnostics10090631

Negi, A., Raj, A. N. J., Nersisson, R., Zhuang, Z., & Murugappan, M. (2020). RDA-UNET-WGAN: An accurate breast ultrasound lesion segmentation using Wasserstein generative adversarial networks. *Arabian Journal for Science and Engineering*, *45*(8), 6399-6410. https://doi.org/10.1007/s13369-020-04480-z

Pacal, İ. (2022). Deep learning approaches for classification of breast cancer in ultrasound (US) images. *Journal of the Institute of Science and Technology*, *12*(4), 1917-1927. https://doi.org/10.21597/jist.1183679

Pang, T., Wong, J. H. D., Ng, W. L., & Chan, C. S. (2021). Semi-supervised GAN-based radiomics model for data augmentation in breast ultrasound mass classification. *Computer Methods and Programs in Biomedicine*, *203*, 106018. https://doi.org/10.1016/j.cmpb.2021.106018

Peng, Y., Tang, W., & Peng, X. (2023). The study of ultrasonography based on deep learning in breast cancer. *Journal of Radiation Research and Applied Sciences*, *16*(4), 100679. https://doi.org/10.1016/j.jrras.2023.100679

Pourasad, Y., Zarouri, E., Salemizadeh Parizi, M., & Salih Mohammed, A. (2021). Presentation of a novel architecture for diagnosis and identifying breast cancer location based on ultrasound images using machine learning. *Diagnostics*, *11*(10), 1870. https://doi.org/10.3390/diagnostics11101870

Qi, X., Zhang, L., Chen, Y., Pi, Y., Chen, Y., Lv, Q., & Yi, Z. (2019). Automated diagnosis of breast ultrasonography images using deep neural networks. *Medical Image Analysis*, *52*, 185-198. https://doi.org/10.1016/j.media.2018.12.006

Vakanski, A., Xian, M., & Freer, P. E. (2020). Attention-enriched deep learning model for breast tumor segmentation in ultrasound images. *Ultrasound in Medicine & Biology*, *46*(10), 2819-2833. https://doi.org/10.1016/j.ultrasmedbio.2020.06.015

Vigil, N., Barry, M., Amini, A., Akhloufi, M., Maldague, X. P. V., Ma, L., Ren, L., & Yousefi, B. (2022). Dual-intended deep learning model for breast cancer diagnosis in ultrasound imaging. *Cancers*, *14*(11), 2663. https://doi.org/10.3390/cancers14112663

Wu, G. G., Zhou, L.-Q., Xu, J.-W., Wang, J.-Y., Wei, Q., Deng, Y.-B., Cui, X.-W., & Dietrich, C. F. (2019). Artificial intelligence in breast ultrasound. *World Journal of Radiology*, *11*(2), 19-26. https://doi.org/10.4329/wjr.v11.i2.19

Wu, T., Sultan, L. R., Tian, J., Cary, T. W., & Sehgal, C. M. (2019). Machine learning for diagnostic ultrasound of triple-negative breast cancer. *Breast Cancer Research and Treatment*, *173*(2), 365-373. https://doi.org/10.1007/s10549-018-4984-7

Xu, Y., Wang, Y., Yuan, J., Cheng, Q., Wang, X., & Carson, P. L. (2019). Medical breast ultrasound image segmentation by machine learning. *Ultrasonics*, *91*, 1-9. https://doi.org/10.1016/j.ultras.2018.07.006

Zhang, E., Seiler, S., Chen, M., Lu, W., & Gu, X. (2019, July 23-27). *Boundary-aware semi-supervised deep learning for breast ultrasound computer-aided diagnosis*. In: Proceedings of the 2019 41st Annual

International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 947-950), Berlin, Germany. IEEE. https://doi.org/10.1109/EMBC.2019.8856539

Zhang, Y., Chen, J.-H., Lin, Y., Chan, S., Zhou, J., Chow, D., Chang, P., Kwong, T., Yeh, D.-C., Wang, X., Parajuli, R., Mehta, R. S., Wang, M., & Su, M.-Y. (2021). Prediction of breast cancer molecular subtypes on DCE-MRI using a convolutional neural network with transfer learning between two centers. *European Radiology*, *31*(4), 2559-2567. https://doi.org/10.1007/s00330-020-07274-x

Zhang, Z., Li, Y., Wu, W., Chen, H., Cheng, L., & Wang, S. (2021). Tumor detection using deep learning method in automated breast ultrasound. *Biomedical Signal Processing and Control*, *68*, 102677. https://doi.org/10.1016/j.bspc.2021.102677

Zhuang, Z., Yang, Z., Raj, A. N. J., Wei, C., Jin, P., & Zhuang, S. (2021). Breast ultrasound tumor image classification using image decomposition and fusion based on adaptive multi-model spatial feature fusion. *Computer Methods and Programs in Biomedicine*, *208*, 106221. https://doi.org/10.1016/j.cmpb.2021.106221