International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Deep Learning Approach for Tumor Segmentation in Breast Cancer

Bhavya Munukurthi*[1], G. V. Gayathri[2]

Submitted:10/03/2024 **Revised**: 25/04/2024 **Accepted**: 02/05/2024

Abstract: In today's world, breast cancer presents a major health challenge, especially for women, necessitating advanced diagnostic methods to enhance patient outcomes. The disease is marked bythe uncontrolled growth of breast cells, forming tumors that can spread if not detected early. To prevent complications from breast cancer, it is crucial to accurately detect and diagnose the condition, followed by providing timely and appropriate treatment. This research explores the use of deep learning to develop a more accurate and efficient system for breast cancer detection. Our methodology starts with the segmentation of ultrasound images using the U-Net algorithm to precisely identify and locate tumors. Post-segmentation, these images are classified using three different models: CNN, VGG19, and MobileNet, to determine if the tumors are benign, malignant,or normal. The findings from our research reveal that the CNN model attained an accuracy of 82%,followed by the MobileNet model reaching 88%, and the VGG19 model outperformed both with an accuracy of 91%. A comparative analysis of algorithms is presented, showcasing the VGG19 model's efficiency in detecting tumors from BUS ultrasound images, in this research paper. This comparison highlights that the VGG19 model delivers the highest accuracy in breast cancer diagnosis among the models evaluated.

Keywords- *Image segmentation, Deep learning, CNN, VGGnet, Mobilenet.*

1. INTRODUCTION

Among the many deadly diseases in the world, breast cancer is one of the most lethal, with high mortality rates. It is a widespread malignancy globally, predominantly affecting individuals assigned female at birth and contributing to a notable mortality rate of 17% [1]. In the spectrum ofdeadly diseases worldwide, breast cancer holds the position of being the second most prevalent cause of mortality among women on a global scale [2]. In breast cancer diagnosis, ultrasound imaging serves as an essential tool for both detection and evaluation. Leveraging sound waves, it generates detailed cross-sectional images of breast tissue. This technique is particularly valuable for further investigating abnormalities identified during mammography or clinical breast examinations, especially in patients with dense breast tissue. Additionally, ultrasound is used in guided breast biopsies [3]. Breast ultrasound can differentiate between fluid-filled cysts and solid masses, and it assists in guiding biopsies. While not as sensitive as mammography or MRI for breast cancer screening, breast ultrasound is valuable for assessing specific areas of concern identified in other imaging tests.

The application of deep learning [4] techniques hold significant promise for improving early detection and diagnosis of breast cancer. This advancement has the potential to translate into moreeffective treatment options by enabling earlier intervention and potentially leading to improved tumor eradication. Image segmentation in breast cancer involves precisely delineating and extracting relevant structures from medical images like mammograms, MRI scans, and ultrasound images. This computational technique aids in diagnosis, treatment planning, and disease monitoring by identifying abnormalities and quantifying tumor characteristics. Deep learning algorithms, capable of analyzing extensive medical image datasets, are revolutionizing breast cancer detection and diagnosis. By automating lesion identification and improving diagnostic accuracy, these models enhance personalized patient care and treatment strategies.

The main contribution of this paper is twofold:

1. Image Segmentation Phase: We introduced a unique approach for addressing breast cancer detection using the U-Net deep learning algorithm to accurately locate tumors in ultrasound images and remove unwanted areas. This crucial step enhances the images, making them suitable inputs for the subsequent classification phase and significantly improving system performance.

2. Classification Phase: Following tumor segmentation, we classify the tumors as either benign ormalignant using three advanced deep learning models: CNN, MobileNet, and the distinctive VGG19 model. These models are applied to the BUSI dataset. To overcome the limited availability of datasets, we utilize data augmentation methods on the BUSI dataset. Our unique contribution lies in thoroughly exploring and implementing deep learning models for classifying breast cancer.

Our proposed work aims to significantly contribute to the advancing realm of breast cancer detection by leveraging cutting-edge technology for enhanced understanding and diagnosis.

The paper following structure is as follows:

Section **2:** This section provides a comprehensive review of existing research related to image segmentation and classification in the context of breast cancer diagnosis.

Section **3:** The proposed methodology is outlined in Section 3. This section includes a detailed description of the deep learning algorithms employed for both image segmentation and classification tasks.

Section 4 contains the results and discussions.

Lastly, Section 5 presents the concluding remarks.

2. Related work

Lately, deep learning methods have shown impressive precision and effectiveness, particularly inearly detection of breast cancer. The integration of image segmentation and classification algorithms, powered by deep learning, has yielded significantly better results and garnered increased attention. By reviewing the existing literature, we aim to provide insights into the methodologies and advancements in image segmentation and classification, thereby contributing to the ongoing development of breast cancer detection.

Wessam M. Salama et al. [5] employed a customized U-Net model for image segmentation and employed various models for classification, including InceptionV3, ResNet50, VGG16, MobileNetV2, and DenseNet121 to analyze mammogram images.

Arevalo et al. [6] utilized mammogram images and employed CNN for feature extraction. They also employed a linear SVM for classifying the data. Their study evaluated the performance and accuracy of CNN-extracted features against manually crafted texture and shape descriptors for mass diagnosis. The findings revealed that the best performance resulted from integrating both learned and manually crafted representations.

Ademola Enitan Ilesanmia et al. [7] introduced an innovative algorithm that employs deep learning to achieve precise tumor segmentation within ultrasound images. Their approach is specifically tailored for segregating breast ultrasound (BUS) images, commencing with resizing the images and subsequently enhancing them via the Contrast Limited Adaptive Histogram Equalization (CLAHE) method to boost image quality. Furthermore, they utilized an enhanced block variation to encode the pre-processed images, leading to an enhanced segmentation performance.

Woo Kyung Moon and team [8] developed a computer-aided diagnosis (CAD) system targeted for tumor detection. Their system utilizes image fusion techniques, incorporating various image representations and ensemble convolutional neural network (CNN) architectures specifically tailored for ultrasound images. Their approach integrates various CNN-based algorithms, including VGGNet, ResNet, and DenseNet. The outcomes highlight the significant impact of diverse image representations on CAD system performance. Specifically, incorporating additional image information enhances prediction accuracy, while integrating tumor shape features enhancesdiagnostic effectiveness.

Marco Caballo and collaborators [9] developed and validated a deep learning network designed for segmenting 2D breast masses in unenhanced breast CT images. They assessed its performance by comparing it with manual annotations made by multiple radiologists. The network was trainedon 58 masses and then tested on 35, covering a total of 93 mass-like lesions. It achieved a Conformity coefficient of 0.85 when compared to an experienced breast radiologist's annotation. Stability and diagnostic accuracy of 672 radiomic features were assessed, revealing that over 90% of the features were stable across different segmentations (ICC > 0.75). MANOVA analyses showed significant results, indicating that the deep learning-based segmentation had similar diagnostic power in distinguishing benign from malignant tumors as expert radiologist annotations.

Siva Teja Kakileti et al. [10] focused on tumor segmentation, a critical component of automated breast cancer analysis. They compared traditional methods like clustering and thresholding with advanced deep learning techniques. Their research highlights the superior performance of various convolutional neural network (CNN) architectures, particularly encoder-decoder models, in termsof accuracy, Dice index, Jaccard index, and inference time, even when working with limited datasets. Infrared imaging, known for its effectiveness in early breast abnormality detection, is often hindered by subjective manual interpretation. Leveraging AI for automated thermal image analysis improves detection accuracy, underscoring its importance for breast cancer screening programs.

Lazaros Tsochatzidis et al. [11] proposed a methodology that enhances CNNs by modifying each convolutional layer to incorporate information from both the input image and its corresponding segmentation map. Furthermore, they proposed a novel loss function that improves upon the standard cross-entropy by incorporating a term that targets the mass region. This novel approach penalizes intense feature activations based on their spatial location. Segmentation maps are either generated from manually annotated ground-truth data or through an automated segmentation process. Experimental findings illustrate that incorporating segmentation information into the CNN enhances diagnostic performance in identifying mammographic breast cancer masses.

Qinghua Huang et al [12] proposed a multi-step strategy for segmenting breast ultrasounds. To begin, a region of interest (ROI) is chosen, followed by image quality enhancement using diverse filters. The cropped image is then divided into superpixels via SLIC, and features are extracted. A bag-of-words model is created from these superpixels for classification using a BPNN. Further refinement is achieved with KNN, leading to the final segmentation. Validation on a BUS dataset comprising 320 cases against five existing methods showcases competitive performance,

particularly in terms of TP and FP rates. The method also delivers precise tumor contour approximations, aligning closely with hand-labelled data across various metrics.

J. Sivamurugan et al. [13] introduce a hybrid model for diagnosing breast cancer, employing optimized deep-learning architecture and benchmark datasets. Image refinement involves Median Filtering, Histogram Equalization, and morphological operations as preprocessing techniques. The segmentation of tumors is conducted using an Optimized U-net-based approach with Adapted- Black Widow Optimization (A-BWO) to ensure precise segmentation. Two detection models are employed: one utilizing GLCM and LGP patterns with DM-OLSTM, and the other employing VGG19, Resnet150, and Inception CNN variants with fused deep features. The final detection output is obtained by averaging scores from both models. This methodology proves to be effective for real-time breast cancer detection applications.

Soner Civilibal et al. [14] employed deep learning techniques to recognize breast tumors in thermal images. They employed Mask R-CNN, a technique that delineates tumors by delineating boundaries around them, simplifying the identification and categorization of both normal and abnormal breast tissues concurrently. By employing transfer learning using ResNet-50 and ResNet-101 models, they improved the precision of tumor detection and segmentation in contrast to earlier investigations. This demonstrates the effectiveness of using a single deep learning model for diagnosing various breast tissue conditions.

3. Methodology

The proposed model involves image segmentation followed by classification of the segmented image data into benign or malignant categories. Figure 1 shows the segmentation process, and Figure 2 shows the classification methods.

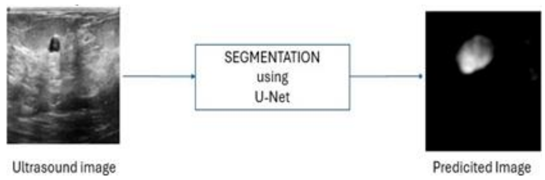

Fig 1. Our proposed segmentation process using U-Net model.

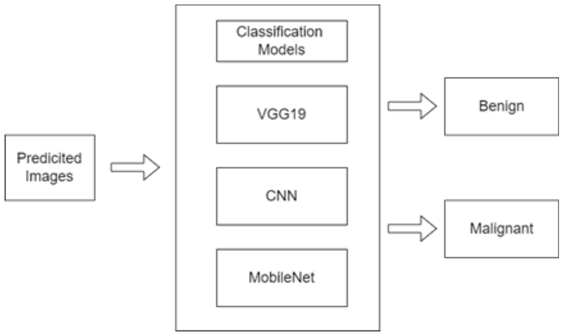

Fig 2. Our proposed classification methods.

3.1 Data Augmentation

We applied data augmentation methods to expand the training dataset using TensorFlow's Keras Image Data Generator. Transformations applied included horizontal flipping, rotation (up to 15 degrees), width and height shifts (within a range of [-10, 10]), and zooming (range of [0.80, 1.00]).These augmentations enhanced the variety of the training data, significantly improving the models' generalization capabilities and overall performance.

3.2 Image segmentation

Our image segmentation strategy leveraged the U-Net architecture, specifically designed for excelling in medical image analysis tasks. U-Net operates by classifying each pixel individually, ensuring the output image maintains the same dimensions as the input. This network is characterized by its two distinct pathways:

Contracting Path (Encoder)

The process begins with the input layer defined to accept grayscale images of size 128x128 pixels. The first convolutional block consists of two convolutional layers, each with 64 filters, a 3x3 kernel size, a ReLU activation function for non-linearity, and the padding is applied to preserve the input dimensions. This block is followed by a max pooling layer with a 2x2 pool size and a stride of 2, reducing the spatial dimensions by half, and a dropout layer with a dropout rate of 0.2 for regularization.

The second convolutional block follows a similar structure but with 128 filters, further abstracting the features. The third and fourth convolutional blocks continue this pattern, each doubling the number of filters to 256 and 512, respectively, while maintaining the dropout rate.

Bottleneck

At the bottleneck of the UNet, two convolutional layers with 1024 filters are applied, maintaining the 3x3 kernel size, and a ReLU activation function. This section captures the most abstract and high-level features of the input image at the smallest spatial dimensions.

Expansive Path:

The expanding path begins with a transposed convolutional layer having 512 filters, a 2x2 kernel size, a stride of 2, no padding is done so the output size is reduced, and ReLU activation to upsample the feature maps. This is followed by a concatenation with the corresponding feature maps from the fourth convolutional block of the contracting path, ensuring that spatial information is preserved. Two convolutional layers with 512 filters and a dropout layer (0.1 dropout rate) follow.

This upsampling and concatenation process continues through the second and third upsampling blocks, with the number of filters halved at each step (256 and 128, respectively). Each block includes transposed convolutions, concatenations with corresponding contracting path feature maps, and dropout layers.

The final upsampling block employs 64 filters, concatenating with the first convolutional block's feature maps, thereby restoring the feature maps to the original input size.

The UNet model's architecture allows it to effectively capture both global and local features through the contracting path, bottleneck, and expanding path. The dropout layers throughout the network help prevent overfitting, making the model robust for image segmentation tasks. This design ensures that the model can precisely assign class labels to each pixel in the input image, making it highly effective for tasks like medical image analysis, where detailed segmentation is crucial.

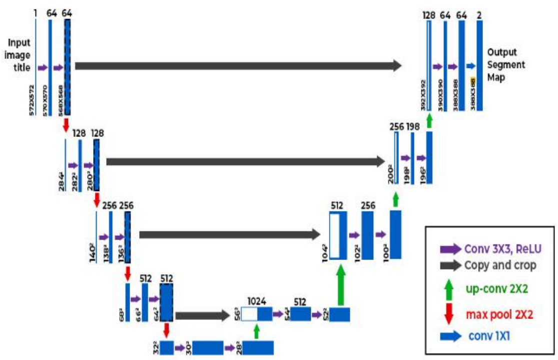

Fig 3: The Basic U-Net Architecture [17]

The network architecture integrates a bottleneck layer that serves as a bridge between the contracting and expansive paths. This layer leverages a series of convolutional operations with a substantial number of filters. This design choice aims to capture intricate details from the highly compressed feature representation at this critical juncture within the U-Net architecture.

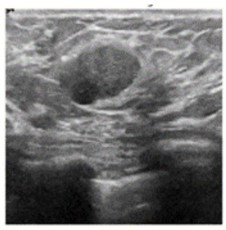

Fig 4. Ultrasound Image

Fig 5. Ground truth

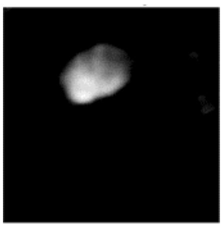

Fig 6. Predicited Image

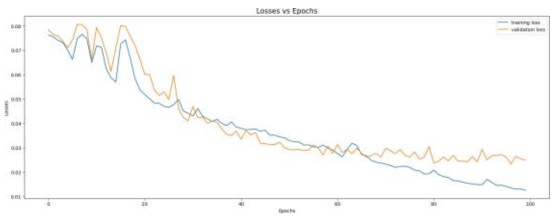

Fig 7. Training and Validation Loss Curve of U-net model.

In Figure 7, the U-Net algorithm demonstrates a robust training process for image segmentation on the BUSI dataset, as depicted in the loss curves. Initially, both training and validation losses start relatively high but gradually decrease over epochs, indicating the model's learning progress. While some fluctuations occur, particularly in the validation loss, the overall downward trend demonstrates the model's efficacy in improving predictions. Although there are signs of potential overfitting in the later phase, the relatively low validation loss suggests decent generalization to unseen data. Overall, the U-Net model exhibits effective training and promising performance for image segmentation tasks on the BUSI dataset.

3.3 Data Splitting

To establish robust model evaluation, we opted for a data-driven approach. The dataset was meticulously divided into two distinct partitions: a training set and a testing set. This separation was achieved using the train_test_split function from the scikit-learn library. To ensure generalizability and prevent overfitting, a 90/10 split ratio was employed for training and testing data, respectively. Furthermore, a fixed random state was introduced to guarantee reproducibility of the split across different experiment runs. Prior to this partitioning process, the data underwent a thorough shuffling procedure to eliminate potential biases and promote representativeness within each subset. This meticulous approach resulted in a training set containing 702 images and a testing set encompassing 78 images. All images were pre-processed and standardized to a uniform size of 128x128 pixels for optimal network performance.

3.4 Model Training

To optimize the training process, the model was meticulously compiled leveraging the Adam optimizer. This optimizer selection aimed to achieve efficient convergence while adapting the learning rate for each parameter individually. A learning rate of 0.00005 was chosen after careful consideration. To assess model performance during training, the mean squared error (MSE) function was employed as the loss function. Furthermore, a ModelCheckpoint callback mechanism was strategically incorporated. This mechanism served the critical purpose of preserving the model iteration that yielded the optimal performance based on the validation loss. The training regime itself spanned 100 epochs, signifying a comprehensive training period. Throughout this process, a batch size of 64 images was utilized to ensure efficient training while managing memory constraints. Notably, both the training and validation datasets were continuously monitored to assess the model's learning progress and prevent overfitting.

During training, the model's architecture involved:

- An input layer for 128x128 grayscale images.

- Multiple convolutional layers with ReLU activation and dropout for regularization.

- Max-pooling layers in the contracting path to downsample the feature maps.

- Transpose convolutional layers in the expansive path are utilized to upsample the feature maps and integrate them with corresponding layers from the contracting path.

- An output layer for segmentation.

The performance of the model was visualized by plotting the training and validation loss over the epochs, showing the progress and convergence of the model during training.

3.5 Classifiers

We employed three different classifiers to evaluate their performance on the ultrasound image dataset: Convolutional Neural Network (CNN), VGG19, and MobileNet. Each classifier was trained and tested on the dataset, and their accuracies were calculated based on their predictions on the test set.

- CNN:

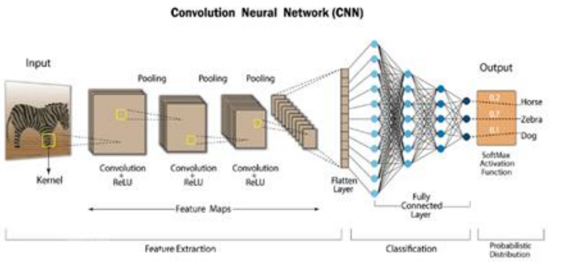

Fig 8. The CNN Architecture. [18]

For image analysis, we employed a powerful Convolutional Neural Network (CNN). CNNs are known for their ability to learn directly from visual data. This specific architecture utilized convolutional layers to identify patterns, followed by batch normalization and dropout for improved performance. Leaky ReLU activation functions were implemented to introduce non-linearity, allowing the model to capture complex relationships. The network culminated in a denselayer for classification, achieving an accuracy of 82%. This success highlights the strength of CNNs in extracting features and spatial information from images, making them ideal for various computer vision tasks like image classification and object detection.

The model starts with an input layer that accepts grayscale images of size 128x128. Four convolutional blocks are added sequentially, each with an increasing number of filters (32, 64, 128, 256) to progressively extract more complex features. The output of these 4 convolutional blocks is flattened into a 1D array. A dense block with 32 hidden units is then added to further process the extracted features. Finally, a dense layer with 3 units and a softmax activation function is included to classify the input images into one of three classes. A LeakyReLU activation function to introduce non-linearity has been used in both the dense layers and the convolutional layer.

- VGG19:

Oxford's Visual Geometry Group introduced VGG19, a renowned 19-layer Convolutional Neural Network (CNN). Debuting in the 2014 paper "Very Deep Convolutional Networks for Large-Scale Image Recognition," VGG19 relies on repetitive 3x3 filters to extract complex features. Its architecture consists of 16 convolutional layers, interspersed with 5 max- pooling layers for dimensionality reduction. Three fully connected layers finalize the network, culminating in a softmax classifier. Prioritizing simplicity with consistent 3x3 filters and 2x2 pooling, VGG19 excels in image classification and computer vision tasks, achieving impressive accuracy on benchmarks like ImageNet (91%).

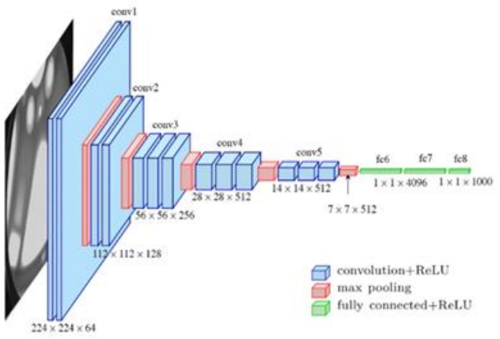

Fig 9. The VGG19 Architecture [19]

In this setup, we leverage a pre-trained VGG19 model, originally trained on the ImageNet dataset, for our breast cancer classification task. The input images are resized to 128x128 pixels with 3 color channels. The classification layers of the existing pre trained model is removed and replaced with an additional set of custom layers to suit our use case. The additional layers consist of the following:

1. A fully connected dense layer with 256 units and a ReLU activation function.

2. A final fully connected layer with a number of units equal to the number of classes, using a softmax activation function for classification.

Before passing the data to these custom layers, the output from the VGG19 convolutional base is flattened into a 1D array. This approach takes advantage of the feature extraction capabilities of the pre-trained VGG19 model and enhances it with custom classification layers suited to our specific task. This transfer learning method enables the model to utilize pre-trained parameters from the ImageNet dataset, improving performance and efficiency on our breast cancer classification dataset.

- MobileNet:

Developed by Google, MobileNet represents a family of Convolutional Neural Networks (CNNs) optimized for mobile and embedded devices. To achieve this efficiency, MobileNet utilizes depthwise separable convolutions, which significantly reduce the number of parameters and computational demands. This focus on efficiency makes MobileNet ideal for tasks where resources are limited, such as image classification, object detection, and segmentation. Despite this focus, MobileNet architectures still achieve impressive accuracy, reaching 88% in this instance.

The model begins by converting grayscale images into RGB images by duplicating the single grayscale channel across three channels, resulting in an RGB image. Next, the pre-trained MobileNet model, originally trained on the ImageNet dataset, is loaded with its top classification layer removed. On top of this base model, a Dense layer with 256 units and ReLU activation is added, followed by another Dense layer with units equal to the number of classes and a softmax activation function for classification. The MobileNet model takes input images of size 128x128, and the output from the base model is flattened into a 1D array before being passed to the Dense layers for final classification.

The accuracies of the models showcase their effectiveness, with VGG19 delivering the best performance, followed by MobileNet and CNN.

3.6 Prediction

After training, each model made predictions on the test set. These predictions were then compared to the true labels to evaluate performance. For CNN, the test images were fed into the model, and the predicted class probabilities were obtained. These probabilities were then converted to class labels using the `argmax` function. The same process was followed for VGG19 and MobileNet, after converting the grayscale images to RGB format as required by these models.

4. Results and Discussion

In our study, we investigated the effectiveness of deep learning models for classifying and segmenting ultrasound images. Specifically, we focused on three classifiers: Convolutional Neural Network (CNN), VGG19, and MobileNet. Our approach involved preprocessing the data, augmenting it, and then training the models to achieve optimal performance.

The custom CNN model demonstrated strong performance with an accuracy of 82%. This notable accuracy can be attributed to the tailored architecture, which included multiple convolutional layers with batch normalization, dropout, and LeakyReLU activations. CNN effectively captured the intricate patterns in the ultrasound images, leading to precise classification results.

Fig 10. Training and Validation Loss Curve of CNN model.

Figure 10 depicts the training process of a CNN on the BUSI dataset for 150 epochs. The x-axis tracks epochs, while the y-axis represents loss values. The training loss (blue) exhibits a steady decline, plateauing at a low value, suggesting successful learning from the training data. The validation loss (orange) also decreases but with fluctuations (spikes) across epochs. Despite these variations, the overall trend is downward, indicating improvement in performance on unseen validation data. However, persistent fluctuations in validation loss might hint at overfitting or noisy validation data. Overall, the CNN effectively learns from BUSI, but further optimization or regularization techniques could improve generalizability.

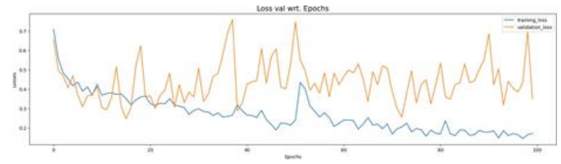

Fig 11. Training and Validation Loss Curve of VGG19 model.

Figure 11 visualizes the training and validation losses of VGG19 on the BUSI dataset for 100 epochs. The x-axis tracks epochs, and the y-axis represents loss. The training loss (blue) exhibits a steady decline, signifying the model's successful adaptation to the training data. In contrast, the validation loss (orange) demonstrates more variability, occasionally mirroring the training loss but with frequent spikes. These fluctuations suggest potential issues like overfitting or noisy validation data, hindering the model's ability to generalize effectively to unseen data.

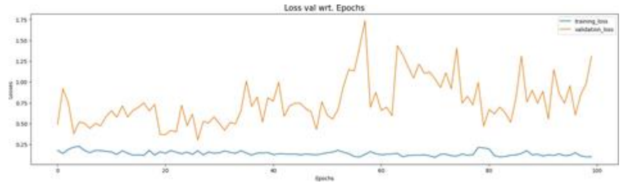

Fig 12. Training and Validation Loss Curve of MobileNet model.

Figure 12 depicts the training and validation losses of MobileNet on BUSI for 100 epochs (x-axis: epochs, y-axis: loss). The training loss (blue) remains consistently low after an initial drop, signifying successful learning from the training data. However, the validation loss (orange) exhibits significant fluctuations with spikes throughout training. This volatility suggests potential overfitting, where the model performs well on training data but struggles on unseen validation data. While MobileNet demonstrates strong training performance, further optimization or regularization techniques might be necessary to improve its generalizability on BUSI.

Table1. Accuracy of CNN model.

	Precision	Recall	F1-score	Support
Benign	0.87	0.79	0.82	42
Normal	0.73	1	0.84	16
Malignant	0.83	0.75	0.79	20
Accuracy			0.82	78
Macro avg.	0.81	0.85	0.82	78
Weighted avg	0.83	0.82	0.82	78

Table 2. Accuracy of VGG19 model.

	Precision	Recall	F1-score	Support
Benign	0.91	0.93	0.92	42
Normal	0.84	1	0.91	16
Malignant	1	0.8	0.89	20
Accuracy			0.91	78
Macro avg.	0.92	0.91	0.91	78
Weighted avg	0.92	0.91	0.91	78

Table 3. Accuracy of MobileNet model.

	Precision	Recall	F1-score	Support
Benign	0.88	0.9	0.89	42
Normal	0.84	1	0.91	16
Malignant	0.94	0.75	0.83	20
Accuracy			0.88	78
Macro avg.	0.89	0.88	0.88	78
Weighted avg	0.89	0.88	0.88	78

The VGG19 model, which was pre-trained on the ImageNet dataset, attained a 91% accuracy following fine-tuning for our particular dataset. Although our custom CNN performed well, the VGG19 model surpassed it in accuracy. This could be due to the complexity and depth of VGG19,which might have been better leveraged despite our relatively smaller dataset.

MobileNet, known for its efficiency and reduced computational requirements, achieved a commendable accuracy of 88%. The model's performance highlights its potential for deployment in real-time and resource-constrained environments, offering a good balance between accuracy and computational efficiency.

5. Conclusion

This study showcases the substantial potential of deep learning, especially CNNs, in classifying and segmenting ultrasound images. The VGG19, achieving an accuracy of 91%, demonstrated strong performance, suggesting its appropriateness for tasks related to detecting breast cancer. Leveraging pre-trained models such as CNN and MobileNet also yielded promising outcomes, with accuracies of 82% and 88%, respectively. These models offer advantages in different scenarios, such as leveraging transfer learning for small datasets (VGG19) and enabling efficient computation in resource-limited settings (MobileNet).Overall, our findings suggest that deep learning models, when properly configured and trained, can significantly enhance the early detection of breast cancer, potentially leading to better patient outcomes. Future work will focus on further optimizing these models, exploring larger and more diverse datasets, and integrating these solutions into clinical workflows for real-world applications.

Conflict of Interest

The author confirm that there is no conflict of interest to declare for this publication.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors would like to thank the editor and anonymous reviewers for their comments that help improve the quality of this work.

References

[1] I. Maniecka-Bryła, M. Bryła, P. Bryła, M. Pikala, ‘‘The burden of premature mortality in Poland analysed with the use of standard expected years of life lost, BMC Public Health 15 (1) (2015) 101. [4] R.A. Hubbard, K

[2] Jose´ Baselga, Mario Campone, Martine Piccart, Howard A Burris III, Hope S Rugo, Tarek Sahmoud, Shinzaburo Noguchi, Michael Gnant, Kathleen I Pritchard, Fabienne Lebrun, et al. Everolimus in postmenopausal hormone-receptor–positive advanced breast cancer. *New England Journal of Medicine*, 366(6):520–529, 2012.

[3] Qinghua Huang, Yonghao Huang, Yaozhong Luo, Feiniu Yuan, and Xuelong Li. Segmentation of breast ultrasound image with semantic classification of superpixels. *Medical Image Analysis*,61:101657, 2020.

[4] Tsochatzidis, Lazaros, Lena Costaridou, and Ioannis Pratikakis. "Deep learning for breast cancer diagnosis from mammograms—a comparative study." *Journal of Imaging* 5, no. 3 (2019):37.

[5] Salama, Wessam M., and Moustafa H. Aly. "Deep learning in mammography images segmentation and classification: Automated CNN approach." *Alexandria Engineering Journal* 60, no. 5 (2021): 4701-4709.

[6] Arevalo, John, Fabio A. González, Raúl Ramos-Pollán, Jose L. Oliveira, and Miguel Angel Guevara Lopez. "Representation learning for mammography mass lesion classification with convolutional neural networks." *Computer methods and programs in biomedicine* 127 (2016): 248-257.

[7] Ilesanmi, Ademola Enitan, Utairat Chaumrattanakul, and Stanislav S. Makhanov. "A method for segmentation of tumors in breast ultrasound images using the variant enhanced deep learning." *Biocybernetics and Biomedical Engineering* 41, no.

2 (2021): 802-818.

[8] Moon, Woo Kyung, Yan-Wei Lee, Hao-Hsiang Ke, Su Hyun Lee, Chiun-Sheng Huang, and Ruey-Feng Chang. "Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks." *Computer methods and programs in biomedicine* 190 (2020): 105361.

[9] Caballo, Marco, Domenico R. Pangallo, Ritse M. Mann, and Ioannis Sechopoulos. "Deep learning-based segmentation of breast masses in dedicated breast CT imaging: Radiomic feature stability between radiologists and artificial intelligence." *Computers in biology and medicine* 118 (2020): 103629.

[10] Kakileti, Siva Teja, Aman Dalmia, and Geetha Manjunath. "Exploring deep learning networks for tumour segmentation in infrared images." *Quantitative InfraRed Thermography Journal* 17, no. 3 (2020): 153-168.

[11] Tsochatzidis, Lazaros, Panagiota Koutla, Lena Costaridou, and Ioannis Pratikakis. "Integrating segmentation information into CNN for breast cancer diagnosis of mammographic masses." *Computer Methods and Programs in Biomedicine* 200 (2021): 105913.

[12] Huang, Qinghua, Yonghao Huang, Yaozhong Luo, Feiniu Yuan, and Xuelong Li. "Segmentation of breast ultrasound image with semantic classification of superpixels." *Medical Image Analysis* 61 (2020): 101657.

[13] Sivamurugan, J., and G. Sureshkumar. "Applying dual models on optimized LSTM with U-net segmentation for breast cancer diagnosis using mammogram images." *Artificial Intelligence in Medicine* 143 (2023): 102626.

[14] Civilibal, Soner, Kerim Kursat Cevik, and Ahmet Bozkurt. "A deep learning approach for automatic detection, segmentation and classification of breast lesions from thermal images." *Expert Systems with Applications* 212 (2023):118774.

[15] Al-Dhabyani W, Gomaa M, Khaled H, Fahmy A. Dataset of breast ultrasound images. Data in Brief. 2020 Feb;28:104863. DOI: 10.1016/j.dib.2019.104863.

[16] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).

[17] Rocha,J., Cunha, A., Mendonça, A.M.:Conventional filtering ver sus U-NET based models for pulmonary nodule segmentation in CT images. J. Med. Syst. 44, 81 (2020). https://doi.org/10.1007/ s10916-020-1541-9.

[18] Shahriar, N. (2023, March 17). *What is Convolutional Neural Network (CNN) - Deep Learning*. Medium. https://nafizshahriar.medium.com/what-is-convolutional-neural-network-cnn-deep-learning-b3921bdd82d5

[19] Bhandare, S. (2023, May 23). *VGG Net architecture explained*. Medium. https://medium.com/@siddheshb008/vgg-net-architecture-explained-71179310050f